Barts and The London
School of Medicine and Dentistry

Identification of genetic factors involved in morphoeic basal cell and sebaceous gland carcinoma of human eyelid tumours with a view to identifying potential treatment targets

John Bladen

MBBS(Hons), BSc, MRCS, PG Cert Clin Lds (Open), FRCOphth

Thesis submitted to the University of London for the degree of Doctor of Philosophy

October 2016

Declaration

I, John Christopher Bladen, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:
Date:

ABSTRACT

Periocular malignancy represents an increasing burden and currently requires disfiguring surgery in an attempt to cure patients. Basal cell carcinoma (BCC) is the commonest cancer worldwide and morphoeic BCC (mBCC) is an aggressive subtype. Sebaceous gland carcinoma (SGC) is a rare, but life-threatening condition that often requires blinding surgery to prevent mortality, especially in the pagetoid subtype.

MBCC has a high risk of local recurrence compared to the more indolent nodular subtype reflected by a different set of driver genes including *FLNB* and *HECTD4*. Surrounding mBCC stroma is abnormal, containing mutations in *EPHA3* and *GLI3*. Four common dysregulated pathways detected using both whole exome and RNA sequencing for mBCC were; 'hedgehog (Hh) signalling pathway', 'BCC', 'Natural killer cell mediated cytotoxicity' and 'Fc Epsilon RI signalling pathway'. Hh mutational profile for nodular BCC was not reflected in the RNA and protein expression. In contrast, Hh overexpression is seen in the tumour and stroma of morphoeic tissue with the latter potentially being partly responsible for its aggressive nature and risk of recurrence that may warrant removal to prevent recurrence.

SGC has a low overall mutational burden, no UV signature and defective mismatch repair signature. Driver genes included *TP53*, *RB1* and the dynein family is a novel driver possibly involved in chromatid segregation as marked chromosomal instability was demonstrated on copy number analysis. Correlation of whole exome and RNA sequencing data demonstrated upregulated 'cell cycle', 'ubiquitin mediated proteolysis' and 'wnt signalling'.

Subtype analysis of pagetoid and nodular SGC revealed the histone gene cluster family as important to both. Oncomir hsa-miR-21 was overexpressed in both and loss of hsa-miR-199a occurs in pagetoid. Increased protein expression of HIST1H2BD was seen in both subtypes as was Hh expression. These novel SGC findings support a chromosomally unstable cancer with the ability to invade extracellular matrix.

DEDICATION

To my Dad, who suffered first-hand the pain of cancer and is now resting in peace. He was thrilled when I was awarded the Fight For Sight Clinical Fellowship and sadly passed away a few months later before I started my endeavours. Some of last words he said to me were make me proud- I hope this is ok Dad and I'll continue to make you proud.

To Sophia and Jacob who highlighted the circle of life and were born in my first year of the PhD. I've been lucky that I could share fully in their first three years and 'you two' never fail to make me smile. Sophia, you are turning out to be savvy girl and already have around 5 secret storage dens around the house. Jacob, you are an astute boy and are picking things up with such speed, including cycling around the house. Both of you have an energy and strength that will carry you to greatness in whatever you pursue in life.

And finally, to the most important person in my life: Dr Mariya Moosajee, a supreme clinician-scientist whose guidance gave me the strength to push through all the obstacles in order to achieve this PhD along with my career to date. Mariya is a loving wife who is the reason why I smile and why I am lucky enough to experience true love, unity and wisdom during the journey of life.

Acknowledgments

I would like to thank Prof Mike Philpott for accepting me into his group and being so supportive by allowing me free rein to develop as a scientist whilst guiding me during the times that I needed help. I look forward to continuing a long-term collaboration and sharing more breakthroughs in the future.

I am indebted to Prof Edel O'Toole who introduced me to Prof Mike Philpott and the Blizard Institute, and opened the doors to allow me to start my journey in research. We actually first met at Kitty's (Dr Catherine Heatley) funeral, so I will thank Kitty too as she was keen researcher who would be pleased with our collaboration - Mariya and I miss her dearly.

I thank Mr George Elia and Miss Emily Austin who went out of their way to accommodate me in preparing my pathological specimens and spending countless hours cutting fresh, frozen specimens in the pursuit of finding useful tumour for my study. Sadly, there were many occasions of up to three hours of cutting without useful tumour, so thank you for your stamina. What is never depicted in a thesis is the countless number of fruitless hours spent to no avail.

I would also like to thank the National Institute Health Research who topped up my funding and Sir Prof Peng Khaw who has been a supportive force working in the background. I would like to thank Mr Dan Ezra who collaborated with the Affymetrix and Nanostring data for the SGC subtype analysis.

A special mention goes to Dr Jun Wang who was keen to collaborate from the outset and took time to explain the world of bioinformatics so that I could learn more quickly and I hope to continue to collaborate for many years to come.

Lastly, and certainly not the least, I would to thank Fight For Sight who funded the PhD and without them I would not have been able to do it. They believed in me which meant a lot, especially as it was a new funding concept for them.

Awards and prizes

Fight For Sight Clinical Fellowship, London 2012

Awarded £160,000

NIHR Eyelid Cancer Genetics Study, NIHR BRC Moorfields Eye Hospital and UCL Institute of Ophthalmology, London 2014

Awarded £18,120

Jean Campbell Cup Pathology Prize, British Association for Ophthalmic Pathology, Glasgow 2015

Identifying genetic factors in periocular morphoeic basal cell carcinoma

Annual Free paper prize, British Oculoplastic Surgery Society, Glasgow 2016

Analysing the nature of periocular infiltrative basal cell carcinoma

Moorfields Alumni Research Medal, London 2017

Morphoeic Basal cell carcinoma: Neo genes and the extracellular matrix

Publications

Abstracts

Bladen JC, Wang JA, Beaconsfield M, O'Toole, EA, Philpott MP Deciphering the code of sebaceous gland carcinoma, RCOPHTH Congress 2016

Bladen JC, Wang JA, Beaconsfield M, O'Toole, EA, Philpott MP. Deciphering the molecular basis of periocular infiltrative basal cell carcinoma Invest. Ophthalmol. Vis. Sci.. 2016; 57(12)

Bladen JC, Wang JA, Beaconsfield M, O'Toole, EA, Philpott MP Identification of genetics factors in periocular morphoeic basal cell carcinoma, Oxford Congress, 2015

Bladen JC, Thaung C, Beaconsfield M, O'Toole, EA, Philpott MP Analysis of the Hedgehog signalling pathway in periocular sebaceous gland carcinoma. Invest. Ophthalmol. Vis. Sci.. 2014, Vol.55, 1282

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 A brief history of cancer

Cancer poses the greatest burden on humanity as the leading cause of global death with figures rising annually not only because of population growth, but also due to our increasing longevity. In 2012, there was an estimated 14 million new cases with 8.2 million deaths from cancer.(Torre et al., 2015) Documentation of cancer has occurred throughout history and recordings have been made during Egyptian, Greek and Syrian times to name but a few (Figure 1.1). The word cancer comes from the Middle English word Canker that denotes a tumour by describing it as a crab or ulcer. It is derived from Latin, which in turn, comes from the Greek word karkinos, meaning a tumour surrounded by veins resembling a crab legs.

Observations recorded by notable physicians hinted to the progressive, temporal nature of cancer by highlighting the different phases of initiation, promotion, and recurrence. Claudius Galen (AD 138-201), a Roman doctor who was taught by a Hippocratist, described liver cancer as a hard, malignant tumour with or without ulceration. The Byzantine physician Paul of Aegina in the early 7^{th} century noted that the uterus and breast were the most frequent site for cancer to form. He goes on to state that surgical excision of uterine cancer is futile due to the high recurrence rate and that breast cancer should be excised without the use of cautery. Although there is no good reason to the refusal of cautery today, excessive tissue destruction can lead to a nidus for infection and may be where the advice stems from. It is certainly used today as electrocautery or surgical diathermy, whereby a high-frequency alternating current causes temperatures of up to 1000°C that results in vaporisation of water within the cell, providing the surgeon with a cutting tool and haemostasis. Moreover, it probably kills any stray cancer cells in its path and so is a beneficial tool for removal of tumours.

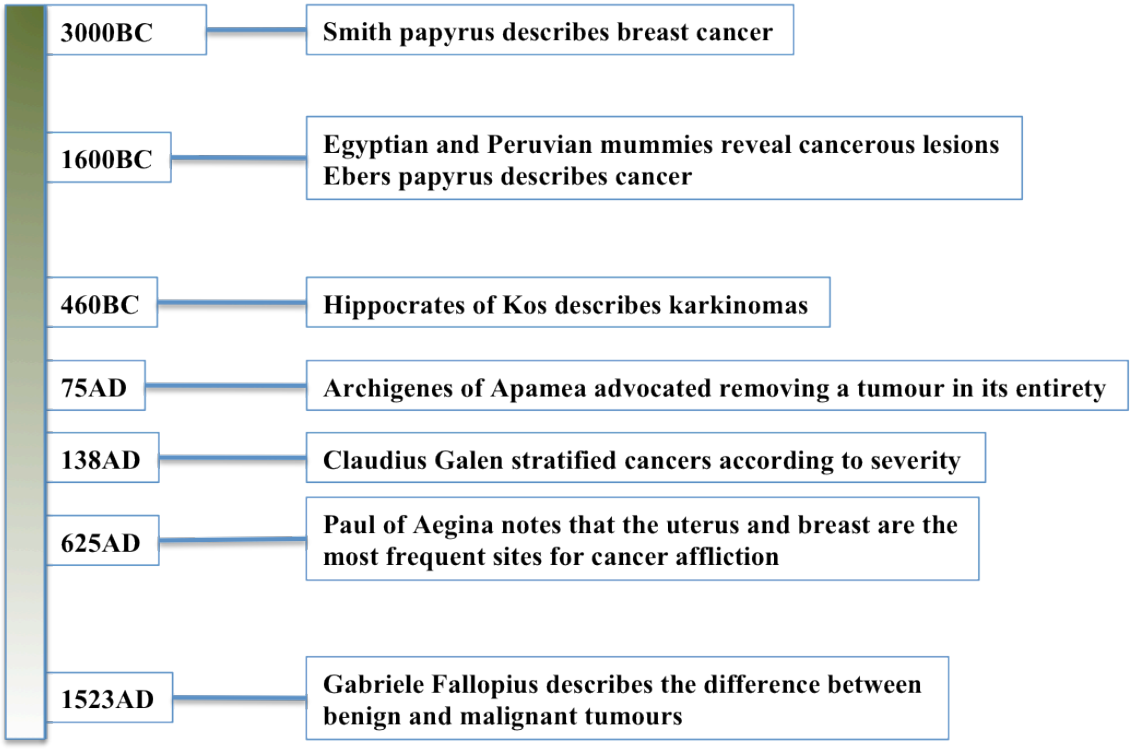


Figure 1.1 Timeline on the historic recordings of cancer.

Visceral cancers were being successfully removed with the improvement in aseptic surgical technique: with the use of gloves, gowns, sterilised instruments and the control of bleeding. A French surgeon Jacques Lisfranc de St Martin resected nine rectal cancers in 1833. Removal of tumours from more audacious parts of the body were attempted with the successful removal of a meningioma in 1888 by the first American brain surgeon, William Williams Keen, Jr. At a similar time, surgeons started to publish the follow up of their patients who had surgical excision of breast tumours.

Cellular descriptions of cancer cells started to become more commonplace at the start of the 20th century with German pathologist David von Hansemann describing the concept of anaplasia in 1902. He observed that tumour cells become more embryonic in appearance and nature with rapid division being the hallmark of anaplasia (Figure 1.2). American pathologist James Ewing collated these descriptions in a book entitled neoplastic diseases.(Ewing, 1919) Furthermore, the German physiologist and Nobel laureate (1931), Otto Warburg, demonstrated the peculiar fermentative type of metabolism seen in the cancer cell. The Warburg effect is the phenomenon whereby cancer cells undergo anaerobic metabolism despite having high levels of oxygen - aerobic

glycolysis.(Warburg, 1924) Subsequently, the enzyme pyruvate kinase muscle (PKM) has been identified as the unregulated culprit in many cancer cells.(Christofk et al., 2008)

Also at the turn of the century, the effects of radiation on neoplastic tissue were investigated. Thor Stenbeck, Swedish röntgenotherapist (early term for a radiologist), used in 1899, fractionated X-rays to treat basal cell carcinoma. Radium (emitting gamma rays) were used at the same time following the discovery by Polish (naturalised French) physicist and Nobel laureate Marie Curie who ironically, died from her life saving treatment: excessive radiation exposure induced aplastic anaemia.

Katsusaburo Yamagiwa, a Japanese pathologist demonstrated that repeated application of coal tar produced skin cancer and so confirmed the presence of environmental risk factors known as carcinogens.(Yamagiwa, 1916) This linked specific chemicals to the aetiology of certain cancers, along with the importance of temporal and dose dependent exposure. Johannes Fibiger, Nobel laureate and Danish pathologist, suggested that cancer models could be developed as shown by the induction of presumed gastric carcinoma in rats fed by cockroaches laced with a nematode called spiroptera (ganglyonema neoplasticum). It was later proven that the changes seen in the rat were merely gastric hyperplasia, however, the concept of developing cancer models was a breakthrough in itself and worthy enough to be awarded the Nobel Prize. In 1932, French pathologist, Antoine Lacassagne, discovered the role of oestrogens in the aetiology of male mice breast cancer by injecting them with folliculin, an observation that greatly stimulated interest in endocrine carcinogenesis. Furthermore, he was the forerunner in the successful treatment of prostate carcinoma with oestrogens (it is still used today in refractory prostate cancer as a salvage treatment). A plethora of carcinogens have since been discovered.

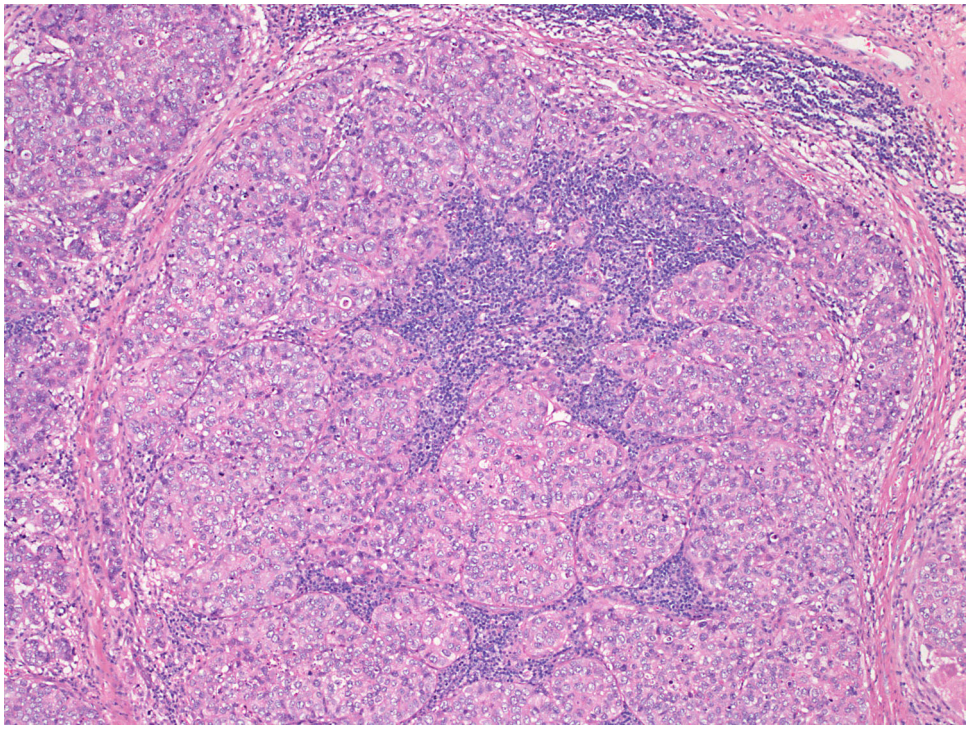


Figure 1.2 Sebaceous gland carcinoma of the eyelid. Rapidly dividing sebaceous gland carcinoma cells that resemble embryonic tissue, an observation made by German pathologist David von Hansemann in 1902 as the hallmark of anaplasia (cancer).

Concerns about cancer led to the formation of many societies, institutes and periodicals devoted to the subject in the early 20th century. Research into cancer exploded and its understanding at a molecular level improved: Coman showed in 1944 that the invasiveness of cancer cells was down to the fact that malignant cells do not stick well to each other compared to their benign counterparts or normal tissue, allowing them to invade locally.(Coman, 1944) Observations made by Murray and Little in 1933 found an extra-chromosomal milk factor caused mammary tumours in mice. This later turned out to be mouse mammary tumour virus, sparking a desperate search for a similar pathogen in humans, but this remains elusive (and a virus may not be the initiator in humans).(Jackson and Little, 1933) We are now aware that a virus can integrate into and modify the hosts DNA, and that cancer is a genetic disorder (see section 1.5). The same group noticed the incidence of mammary cancer in hybrid mice was determined by the degree of cancer susceptibility of the mother highlighting the effect of inherited factors on cancer development. Many inherited cancers have subsequently been identified such as retinoblastoma, however, even if we don't directly inherit a cancer forming gene, our genetic makeup will influence the risk of developing a particular type of cancer, making certain individuals more susceptible when exposed to an identical environment factor.

Modifying these genetic risk factors before the cancer develops could involve having augmented viruses inserted into our genome that corrects or removes the defective gene. Gene therapy attempts to remove these defects and up until recently relied on reprogramming viruses to insert the correct gene amongst other strategies.(Freytag et al., 2007) In fact, there is a whole journal aptly named Cancer Gene Therapy dedicated to this task. Breakthrough technology using clustered regularly interspaced short palindromic repeats (CRISPR)-associated enzyme Cas9 (derived from bacteria) to amend (insert and remove) DNA sequences with great accuracy, could be used for this process in the future.(Sternberg et al., 2014)

1.1.1 Basic language of cancer today

Tissues normally divide when new cells are required, however, loss of their normal constraints for reproduction lead to an expanding clone of abnormal cells: cancer. If this remains localised, it is described as benign, whereas invasion or colonization of territories that are normally reserved for other cells, is described as malignant. Some cancers follow a stepwise transition from benign, premalignant to malignant, however, this is certainly not a universal phenomenon. There may be a stepwise transition at a molecular level whereby the cell gains more mutations and diverges away from normality, but a simple linear transition is sadly far from reality. Malignancy is a spectrum, with some being slow to grow or invade, giving time for their detection and treatment versus aggressive, locally invasive tumour that metastasis early. What determines the cancers ability to locally invade, spread and plunder the body's cellular machinery to its own advantage is the subject of this research and this dearth of data in eyelid cancer requires attention.

1.1.2 Non-melanoma skin cancer –the world's most common cancer

Non-melanoma skin cancers (NMSC) are the most common human cancers and their incidence continues to rise.(Madan et al., 2010, Kwasniak and Garcia-Zuazaga, 2011) The vast majority of eyelid NMSC are basal cell carcinomas (BCC) with 5% attributed to sebaceous gland carcinoma (SGC).(Shields et al., 2005) Morphoeic BCC (mBCC) is a particularly aggressive subtype, which invades local tissue, such as the eye socket and adjacent structures including the nose, in an uncompromising manner. SGC is a masquerader of benign conditions; it is often picked up late and results in death in up to

22% of patients.(Boniuk and Zimmerman, 1968, Doxanas and Green, 1984, Rao et al., 1982) Both can lead to significant eye morbidity from the disease itself along with the tissue destructive surgical treatment. In addition, some cases require total removal of the eye and its supporting structures (exenteration) leading to blindness. BCC numbers are increasing by 10% worldwide per annum and will soon be more prevalent than all other cancers combined.(Madan et al., 2010) A growing older population and amplified UV light exposure may in part explain this increase, but they are unlikely to be the only factors. MBCC and SGC require more invasive treatment modalities and little is known on the development of these cancers. This lack of knowledge combined with their increasing prevalence demands detailed research into the behaviour of these eyelid tumours.

1.2 Periocular cancer

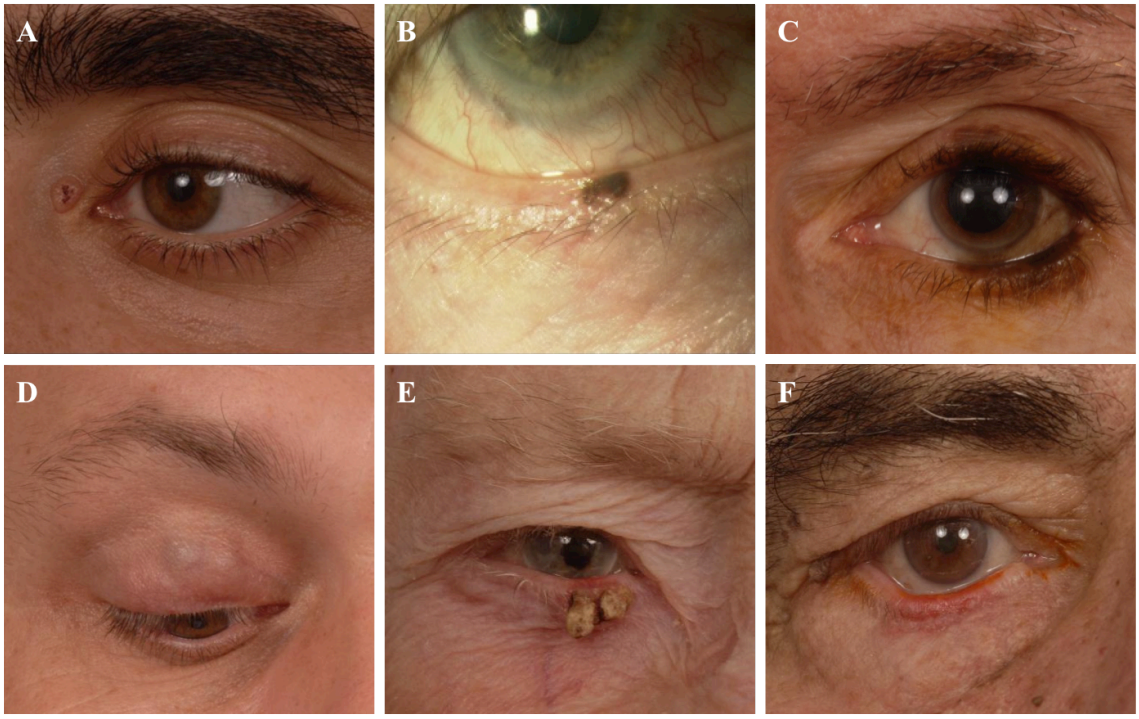


Figure 1.3 Periocular malignancies: (A) **Trichoepithelioma at the medial canthus.** This mimics nodular BCC, but is a benign neoplasm of the hair follicle. (B) **Melanoma in-situ.** A well circumscribed lesion that could easily be labelled as a naevus, however, was actually a melanoma in-situ. (C) **Lentigo maligna** of the upper and lower eyelid. This is a melanoma in-situ that can undergo invasive change to melanoma proper. (D) **Invasive Melanoma of the upper eyelid**. A discrete mass within the upper eyelid and not obviously on the skin surface, which could be misdiagnosed as a chalazion, however, the dark pigmentation gives a clue to its true nature (E) **Actinic keratosis with a keratin horn of the lower eyelid**. A premalignant condition that can transform into squamous cell carcinoma. It can also present as a flat scaly lesion. (F) **Squamous cell carcinoma of the lower eyelid.** A plaque type of SCC represents around 2-5% of eyelid malignancies, but has a greater propensity to spread compared to BCC.

1.2.1 Local anatomy and associated tumours

The periocular region (defined as the area within the orbital margin surrounding the globe) encompasses many different tissue types and comprised of orbital content (extraocular muscles, fat, optic nerve, cranial nerves and blood vessels), conjunctiva, eyelids (skin, tarsal plate and orbicularis oculi muscle), lacrimal system (tear production and drainage system) and the surrounding bone. An aberration can occur in any of these tissues and so a wide variety of tumour is seen in a very small space (Figure 1.3). Moreover, a tumour can arise from the globe (extraocular extension) and mimic a periorbital lesion. Each tissue or compartment has its own commonest primary cancer and are listed in Table 1.1. This makes diagnosis more challenging and stresses the importance of obtaining a histological sample in order to characterise the tumour type; even on a microscopic level it can be a balance of probability with the clinical details being a decisive factor (such as speed of onset, age, associated symptoms to name a few). Choroidal melanoma is a cancer where a histological biopsy is avoided as the diagnosis can often be made on clinical grounds alone; a biopsy has to traverse the globe structures and has the potential to be either sight damaging or seed (spread) the cancer. Lacrimal gland (pleomorphic) adenoma is another tumour where a biopsy is avoided. Not because there are delicate structures to traverse with the biopsy needle, rather that piercing of the tumour capsule causes seeding and an unacceptable recurrence rate. Fortunately, classical features of lacrimal fossa expansion, indentation of the globe, well-circumscribed lesion with calcification present in 30%, allows for a radiological diagnosis and subsequent treatment with surgical removal of the intact tumour with a surrounding pseudo-capsule to prevent seeding. Although all tumours are encompassed in the term periocular, eyelid skin and conjunctiva are focused on below.

Compartment *Possible tissue of origin*	**Malignant**	**Premalignant**	**Benign**	**Comment**
Eyelid				Commonest
Hair follicle	Basal cell carcinoma			Commonest
Hair follicle	Squamous cell carcinoma	Actinic keratosis. Bowens disease		
			Squamous cell papilloma	
Hair follicle	Keratocanthoma			
Dendritic cells (neural crest)	melanoma	Lentigo maligna		
Epidermal stem cells	Merkel cell carcinoma			
Human herpes simplex virus 8 infected endothelial cells	Kaposi's sarcoma			
	Primary eccrine porocarcinoma Apocine carcinoma Sebaceous gland carcinoma			
Conjunctiva	Squamous cell carcinoma	Carcinoma in-situ (CIN)		
Orbit				
Optic nerve	Glioma			
Blood vessels	Angioma			
Connective tissue	Rhabdomyosarcoma			Commonest malignant orbital lesion in children
Smooth muscle	Leiomyoma			
Lymph	lymphoma			
Lacrimal gland	Adenocystic carcinoma Mucoepitheliod Pleomorphic adenocarcinoma		Pleomorphic adenoma	
Bone	Osteosarcoma		Osteoma	
Globe				
Retinal	Retinoblastoma			Commonest ocular tumour in children
	Astrocytoma			
Uveal *Melanocyte (neural crest cells)*	Uveal melanoma (iris, ciliary body, choroidal)			

Table 1.1 Periocular tumours grouped by compartment. Each compartment has its own commonest primary cancer.

1.2.2 Periocular skin and conjunctiva

The eyelids are essential for maintaining corneal health, which in turn, affords clear vision. Their role is to maintain the wetting of the cornea with lacrimal secretions using a wiper effect, contribute to tear production and drainage, protect the globe from injury (blink reflex) and modulate the light entering the eye including obstructing it completely during sleep. The eyelid is often split into anterior and posterior lamellae for descriptive

purposes (Figure 1.4). This stems from the fact that during development, folds cover the cornea in two phases (during the seventh week and seventh month of embryogenesis) separated by a mesodermal layer that differentiates into the tarsus and orbital septum. Surgically, the two layers can be split at the site of the grey line and serves as an important dissection plane where minimal scarring occurs. The anterior lamella comprises of skin, subcutaneous tissue, and muscle fibres of the orbicularis oculi. The posterior lamella comprises orbital septum/tarsal plate, smooth muscle and palpebral conjunctiva.

The skin of the eyelid is unique in that it is the thinnest in the body as it contains no fat and undergoes the most movement (you blink 16 times a minute). It is fairly elastic and forms many folds, most notably the superior palpebral furrow of the upper eyelid. The skin is stratified squamous containing numerous melanocytes, with less sebaceous glands, sweat glands, hair follicles and dermal papillae compared to normal skin with the exception of the eyelid margin where they are numerous. Below, the subcutaneous tissue contains abundant elastic tissue. The hairs at the margin are coarser forming eyelashes, the connective tissue is denser and more collagenous, and the dermal papillae are higher. The eyelashes are arranged in two or three rows with 150 in the upper eyelid and 75 in the lower. They are replaced every 4 to 6 months, but can grow back within 2 if cut. A modified sebaceous gland of Zeiss opens into each hair follicle and between them are the eccrine sweat glands, sebaceous glands and apocrine sweat glands of Moll. The grey line at the eyelid margin depicts the junction between the skin and the non-keratinised conjunctival epithelium, just in front of the orifices of the meibomian glands which reside in the tarsal plate. The meibomian is a large holocrine (sebaceous) gland, secreting a phospholipid that prevents the evaporation of the aqueous portion of the tear film and increases surface tension. The skin lymphatics drain to the anterior cervical chain inferomedially, submandibular glands inferiorly and the parotid gland laterally. Tumours that invade the lymphatics such as sebaceous gland carcinoma (SGC) typically spread to these nodes first and are often sampled when extensive disease is treated by exenteration.

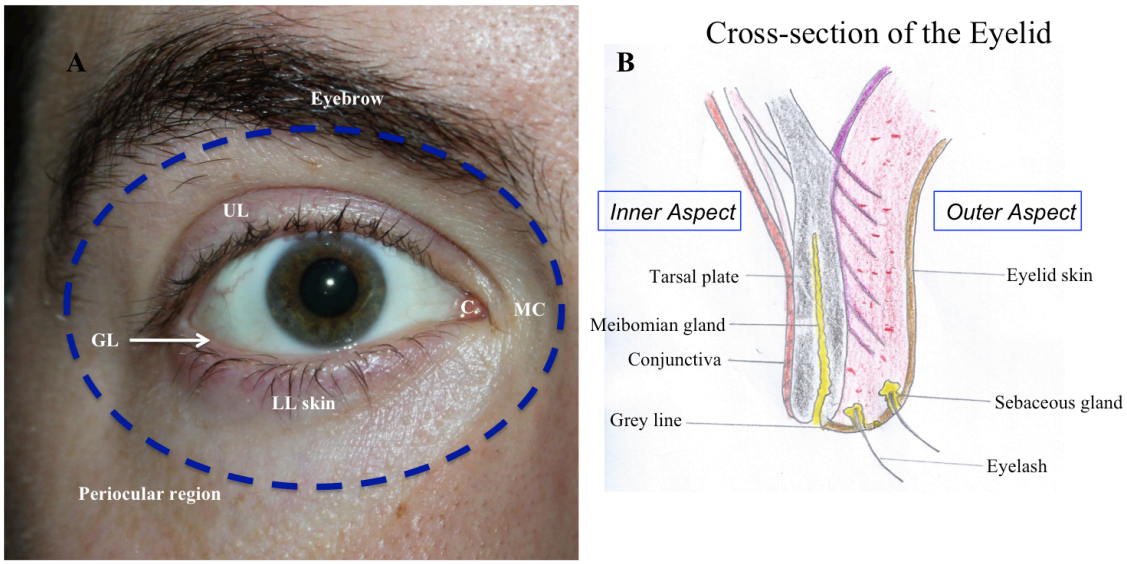


Figure 1.4 Periocular region. (A) Photograph of the periocular region that is contained with the orbital margin denoted by the blue dotted line. (B) A cross section of the eyelid demonstrating the internal structures and the site of the grey line that separates it into anterior and posterior lamellae. UL, upper eyelid; LL, lower eyelid; GL, grey line; C, caruncle; MC, medial canthus.

The conjunctiva is 2-5 layered stratified squamous (non-keratinised) at the eyelid margin, thins out to 2 layers with a cuboidal structure over the tarsal plate of the posterior lamella (palpebral conjunctiva) and changes into stratified columnar over the globe (bulbar conjunctiva), only to return to a squamous structure at the corneal limbus. Scattered within the epithelium are mucin secreting goblet cells and microvilli secreting glycoproteins. The epithelium sits on a stromal nest (substantia propria) containing a superficial lymphoid and deep fibrous layer. The former is essential for immune surveillance and contains mast cells, lymphocytes, plasma cells and neutrophils. In addition, there are accessory lacrimal glands of Krause and Wolfing. The fibrous layer contains the neurovascular complexes and lymph vessels; it almost thins to non-existence over the tarsal plate where the conjunctiva is very adherent to the tarsal plate. At the inner part of the eye a semilunar fold (plica semilunaris) of bulbar conjunctiva is formed surrounding a bulge known as the caruncle which contains large sweat gland and lanugo hair. The plica semilunaris probably represents a rudimentary nictitating membrane seen in animals. The plica stroma can contain cartilaginous tissue and the caruncle stroma can contain striated muscle fibres.

Periocular carcinomas, especially BCCs, have a predilection for the eyelid margin with their excision being potentially hazardous for future eyeball protection. Excision of the

lesion if fairly straight forward as long as you avoid damaging the globe. More difficult is the reconstruction of the eyelid post excision: it is essential to try and regain some of the aforementioned eyelid properties and the type of reconstruction depends on the amount of skin lost during removal. In addition, both the anterior and posterior lamellae need to be replaced ideally. Loss of less than a third can be repaired with direct closure, between a third to a half using a modified Beard-Cutler or Hughes flap for the lower and upper eyelid respectively. Removal of over half of the eyelid warrants rotational myocutaneous flaps remembering to replace the posterior lamella with a mucous membrane graft or equivalent.

1.3 Basal cell carcinoma

BCCs are the commonest cancers in the world with an incidence of 100 in 100,000 in the UK and highest in Australia at 884 in 100,000.(Staples et al., 2006, Madan et al., 2010) This suggests that there is a strong environmental risk factor, namely UV light exposure (both solar irradiance and sun beds). The head is a common location for BCC supporting the UV theory, however, up to a third of BCC occur in areas with minimal or no UV light exposure.(Firnhaber, 2012) Our increased longevity makes age a factor for developing BCCs, although the incidence is increasing in the young too.(Kwasniak and Garcia-Zuazaga, 2011) Other environmental risk factors include arsenic, coal tar, ionizing radiation demonstrated by increased incidence in Nagasaki, and immunosuppression.(Guo et al., 2001, Kishikawa et al., 2005, Berg and Otley, 2002) Genetic factors are highlighted by the susceptibility to BCC according to skin type. Fitzpatrick type 1 (light skin, eyes and hair colour) has the highest risk of UV-induced skin lesions including BCC.(Ravnbak, 2010) Specific genetic changes have been eluded too from studying basal cell naevus syndrome (BCNS aka Gorlin syndrome) and are discussed in section 1.5.

1.3.1 Macroscopic features

High risk macroscopic factors include mid facial location (H-zone, see Fig 1.6) large size, longstanding and recurrence. Local destruction is where the complications occur, making the periocular region particularly vulnerable (Figure 1.5). The term ‘rodent ulcer’ was

given to BCC, as left untreated, the rodent ulcer will slowly ‘eat’ the skin coverings exposing the underlying flesh and bone.

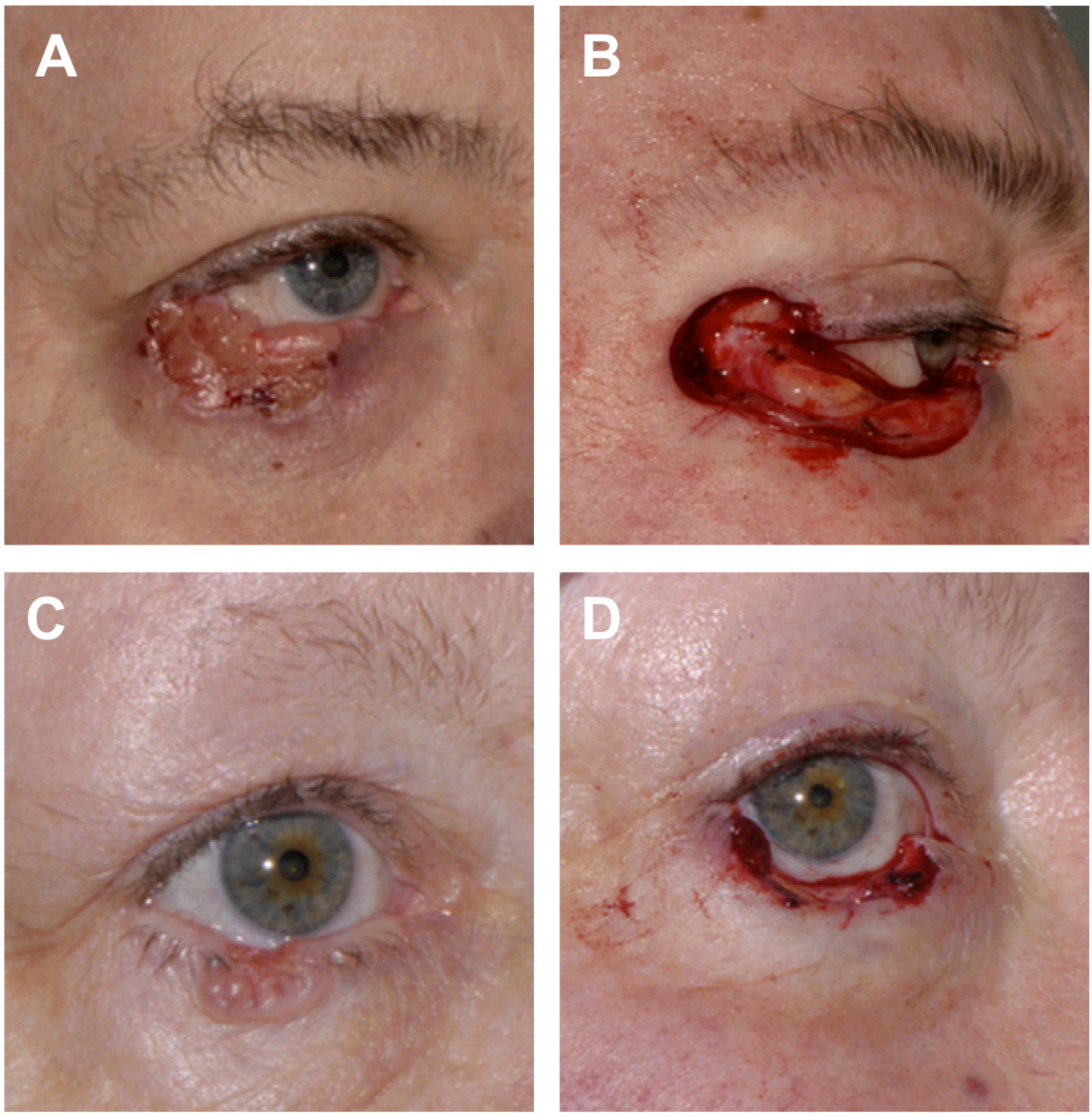


Figure 1.5 Macroscopic appearance of basal cell carcinoma. (A) Morphoeic BCC of the lower eyelid with loss of normal structures and ulceration of the skin. (B) Post-Moh’s excision of morphoeic lesion demonstrating loss of the lower and the lateral canthus in its entirety along with a portion of the upper eyelid. (C) Nodular BCC of the lower eyelid with notching of the eyelid margin, ulceration and distortion of normal structures plus loss of eyelashes. (D) Post-Moh’s excision of the nodular lesion demonstrating loss of all the lower eyelid.

Figure 1.6 High risk H-zone. The H-zone consists of the periocular, nasal alar, nasolabial fold and auricular regions with confers a higher risk of BCC recurrence. A large basal cell carcinoma is seen on the right lower eyelid whose histology revealed a morphoeic subtype and also confers a high risk of recurrence.

1.3.2 Microscopic features

There are several different classifications for BCC, however, a clinically relevant description relating to patient risk is laid out in the RCpath Dataset.(Slater, 2014) Growth pattern and differentiation are the two microscopic factors that relate to risk. Growth pattern is divided into low and high risk subtypes. The former subtypes are superficial, nodular and fibroepithelial. The nature of BCCs are classically associated with these subtypes as they are indolent in nature. Superficial BCC consists of a focal collection of germ cells in close contact with the epidermis or a hair follicle, does not extend beyond the papillary dermis, is less than 1mm thick and often coupled with a stromal reaction. It is not clear if superficial BCC represents a form of in-situ disease, but this may explain its localised nature making it readily amenable to topical immunomodulatory treatment.

Nodular BCC (nodBCC) are round lesions greater than 0.15 mm in diameter (differentiating them from micronodular; a high-risk subtype). They can be cystic, pseudoglandular or rippled in appearance (Figure 1.6). Fibroepithelial are small pedunculated lesions and can be regarded as a benign trichblastoma, but the WHO consider it a BCC.

High risk subtypes include morphoeic (as known as infiltrating or sclerosing) and micronodular. MBCC contain irregular, spiky strands of cells often accompanied by stromal fibrosis (Figure 1.7). Micronodular are small round lesions less than 0.15 mm, often with less than 25 cells per island. As mentioned, the individual growth patterns are fairly academic and it has been proposed that the high-risk samples should be grouped under one umbrella term: infiltrative.

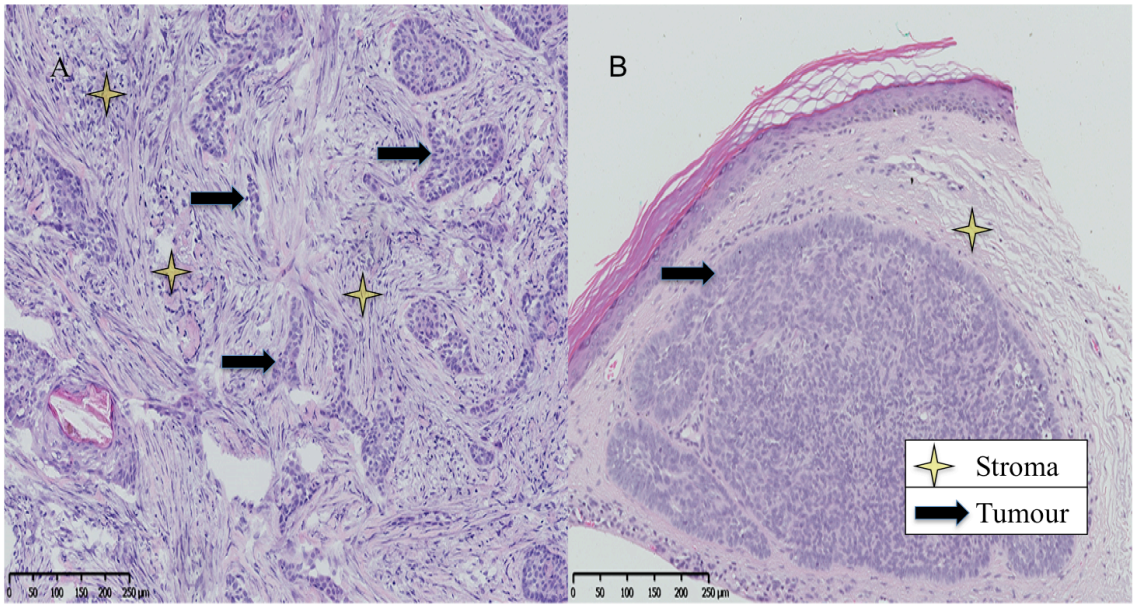


Figure 1.7 Microscopic appearance of basal cell carcinoma. (A) Morphoeic BCC with irregular, spiky strands of tumour cells. (B) Nodular BCC with a well circumscribed collection of cells with palisading of the outer cells.

Microscopic heterogeneity exists, with samples often a mixture of the above subtypes. Any sample that contains both high and low risk (Table 1.3) subtypes must be defined as high risk, no matter how small the proportion of the high risk component.(Network, 2013)

There are a vast number of different histological differentiation patterns including pigmented, clear cell, eccrine, follicular and sebaceous to name a few. These represent no increased risk, however, squamous (basosquamous) differentiation with moderate or severe atypia is associated with metastasis and local recurrence.(Le Boit PE, 2008)

Staging of the cancer is done using the tumour, node and metastasis (TNM) categorisation of malignant tumours (Table 1.2).(Sobin LH, 2009) Perineural and lymphovascular invasion are high risk factors with the former a determinant of tumour staging T4. The latter has less robust evidence to be a T4 determinant and the risk of metastasis or recurrence is less than perineural invasion, but is usually a feature if basosquamous BCC; in this context, it does determine an increased risk of metastasis. Lymphovascular invasion without squamous differentiated is noted in the pathology report, although currently its relevance is unknown.

Tumour		Node		Metastasis	
Symbol	**Description**	**Symbol**	**Description**	**Symbol**	**Description**
Tx	Tumour not assessable	Nx	Regional lymph node not assessable	M0	No distant metastasis
Tis	Carcinoma in-situ	N0	No lymph node involvement	M1	Distant metastasis
T0	No primary tumour	N1	Single, ipsilateral lymph node ≤30 mm size		
T1	≤ 20 mm	N2a	Single, ipsilateral lymph node, >30 mm to ≤60 mm		
T2	≥ 20 mm	N2b	Multiple, ipsilateral lymph nodes ≤60 mm		
T3	Local invasion of bone (orbit, temporal or maxilla)	N2c	Bilateral or contralateral lymph nodes ≤60 mm		
T4	Local invasion of skeleton (axial or appendicular) or skull base or perineural invasion	N3	Lymph node ≥ 60mm		

Table 1.2 TNM staging adapted for BCC. T, tumour; N, node; M, metastasis

Low risk clinical features
>24yr history
No immunosuppression
No genetic predisposition
Below the clavicle
<10 mm in size
Not recurrent
Not persistent
Not infiltrating
Not on an important structure
Not highly visible with an either a poor cosmetic risk or difficult closure

Table 1.3 Low risk BCC features. BCC that meet these clinical features could be managed in the community. Adapted from NICE guidelines.(National-Cancer-Peer-Review-Programme, 2014)

1.3.3 Management

There are a vast number of treatment modalities for BCC which reflects the tumour heterogeneity (i.e. the different nature of the subtypes mentioned earlier) and the lack of an ideal treatment for this condition. Surgery has been the mainstay of treatment and still is when it comes to the removal of high risk subtypes (Table 1.4). In such cases, Mohs micrographic surgery is employed to ensure that all of the cancer is cleared in one sitting. It involves real time histology, i.e. assessment of the tumour under the microscope during the surgical removal. It is labour intensive, requires a large team of staff to rapidly process the tissue and can last up to 8 hours, making it stressful for patients. And this excludes reconstruction which tends to be done the following day. Nevertheless, Mohs micrographic surgery is the gold standard for high-risk cases. Standard surgical excision normally involves removal with a 3mm margin of tissue, although this varies between 3-15mm.(Telfer et al., 2008)

Status	Description
Growth pattern	Infiltrative
Differentiation	Basosquamous
Level of invasion	Deeper than subcutaneous fat
Perineural invasion	Present
Lymphovascular invasion in basosquamous	Present
TNM staging	pT2, T3, T4
Margins	Involved or <1 mm clearance

Table 1.4 High risk BCC features. Summary of high risk microscopic features. Adapted from the Royal College of Pathologists Dataset 2014 (Slater, 2014)

1.3.4 Spread and Prognosis

The risk of metastasis is generally low (0.2%) if you include both low and high risk BCCs together. Due to small numbers of high risk BCCs, it is difficult to have a reliable outcome, but they have been quoted as having 2% risk of metastasis if looked at independently.(Snow et al., 1994) Five year recurrence rates depends on treatment modality (Table 1.5).

Treatment	RTR %	5yDFR %	Reference
Surgical excision low risk 3mm	15		(Breuninger and Dietz, 1991)
Surgical excision low risk 4-5mm	5		(Kimyai-Asadi et al., 2005)
Surgical excision high risk 3mm	34		(Breuninger and Dietz, 1991)
Surgical excision high risk 5mm	18		(Breuninger and Dietz, 1991)
Surgical excision high risk 15mm	5		(Breuninger and Dietz, 1991)
Mohs mSx for primary BCC		98.6-100	(Malhotra et al., 2004)
Mohs mSx for recurrent BCC		92.2-96	(Leibovitch et al., 2005)
Curettage and electrodissection	33		(Suhge d'Aubermont and Bennett, 1984)
Radiotherapy for primary BCC		91.3	(Rowe et al., 1989a)
Radiotherapy for recurrent BCC		90.2	(Rowe et al., 1989b)
PDT for superficial and nodular BCC	25		(Wang et al., 2001)
Cryotherapy periocular low risk cases		8	(Buschmann, 2002)
5% Imiquimod in superficial BCC	18		(Geisse et al., 2004)
SMO inhibitor (Vismodegib) for *BCC	50-66		(Gill et al., 2013, Sekulic et al., 2015)

Table 1.5 BCC treatments. List of various modalities used to treat BCC highlighting the percentage of patients with residual tumour using that technique and risk of recurrence by showing the percentage of patients who are disease free after 5 years of treatment. RTR %, percentage residual tumour rate; 5yDFR %, percentage of 5-year disease free rate; mSx, micrographic surgery; *BCC not amenable to other treatment modalities

1.4 Sebaceous gland carcinoma

Sebaceous gland carcinoma (aka sebaceous cell carcinoma, cutaneous sebaceous carcinoma, sebaceous adenocarcinoma) is a rare aggressive cancer that has a predilection for the periorbital region, perhaps due to the multitude of glands surrounding the eyeball, but can come from extraocular sites, albeit mainly within the head region.(Deprez and Uffer, 2009) The incidence in the UK is estimated at 1 in 1000,000. Geographical variation is significant with the incidence around 0.65 per 100,000 in Canada versus China where it represents almost a third of the malignant eyelid workload and second to

BCC in frequency.(Kuzel et al., 2012, Xu et al., 2008) In a Japan, the rate of SGC equalled that of BCC in one study.(Obata et al., 2005) Risk factors include radiation exposure such external beam radiotherapy which used to be the mainstay of treatment for retinoblastoma (Kivela et al., 2001). Immunosuppression with HIV or the presence of HPV may also be a risk factor.

1.4.1 Macroscopic features

There are two broad macroscopic presentations, namely nodular (local) or pagetoid (spreading). Misdiagnosis is common, often labelled as a benign chronic meibomian cyst (chalazion) in its nodular form or conjunctivitis in its conjunctival pagetoid type (Figure 1.8). Thus, it is described as a masquerader lesion with delays to diagnosis often greater than a year. If it involves the eyelid margin, madarosis (loss of eyelashes), atrophy, thickening or notching of the margin is seen. Pagetoid expansion of the skin or conjunctiva (conjunctival intraepithelial invasion) carries a higher risk of removing the whole of the orbital content (exenteration). (Chao et al., 2001) In contrast to BCC, SGC more commonly effect the upper eyelid and the majority are unicentric.(Rao et al., 1982, Ni, 1982) When cut, the nodular form can be yellow-white or ivory in colour reflecting the fatty content that is associated with SGC.

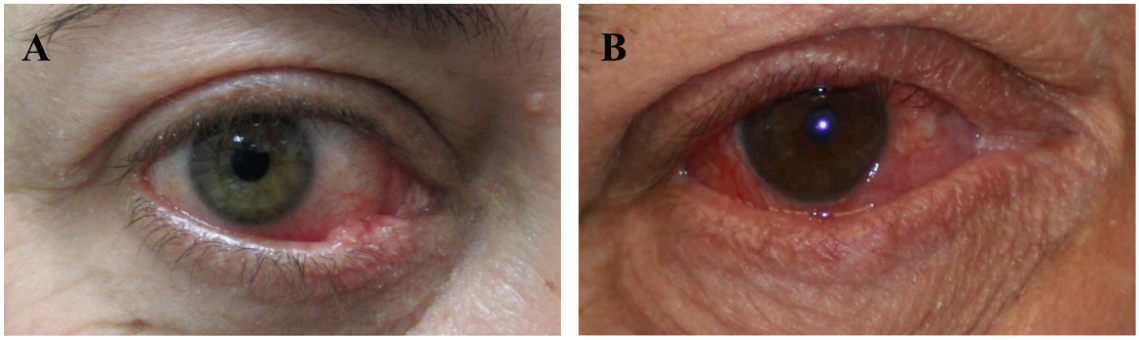


Figure 1.8 Macroscopic features of SGC. (A) Nodular (local) SGC presenting as an eyelid lump. (B) Pagetoid (spreading) SGC presenting as a red eye and misdiagnosed as conjunctivitis for 18 months.

1.4.2 Microscopic features

The presumed cell of origin is the holocrine sebaceous gland that is distributed throughout the skin except the palms of the hand and soles of the feet. They are found within the tarsal plate as meibomian glands, the caruncle, at the mucocutaneous junction and when

associated with the eyelashes, are termed glands of Zeiss. Nevertheless, there is marked morphological variability in terms of cytology and architecture amongst SGC tumours. Compounded by the fact that it is rare, there is no universal classification and diagnosis can be difficult if the 'classical features' are not present.

'Classical features' include positive staining for intracytoplasmic fat in frothy acinar-like arranged cells (Figure 1.9). This make the diagnosis fairly straightforward in well differentiated tumours. Perilipin 2 (PLIN2) stained positively in the majority of a small case series of 9: the group found that a panel of stains (PLIN2; cyclin-dependent kinase Inhibitor 2A, CDKN2A; mucin 1 cell surface associated, MUC1; androgen receptor (AR), keratin 7, type II, KRT7; marker of proliferation Ki67, MKI67) were useful in differentiating SGC from BCC or squamous cell carcinoma (SCC).(Ali-Ridha, 2014) SGC stained four antibodies: MUC1, PLIN2, CDKN2A and AR; BCC exhibited positive staining for KRT7; SCC stained for MUC1 only. A similar small case series supported the use of PLIN2 for SGC diagnosis and the presence strong AR staining in pagetoid spreading type, although the latter was noted in varying intensities in both BCC and SCC. (Becker, 2014)

Staging involves the TNM system and similar to Table 1.3. Microscopic features can be divided into growth pattern and differentiation, but in contrast to BCC, it they have a different meaning (Figure 1.9). In SGC, growth pattern relates to the lineage of differentiation, which essentially means sebaceous cell (rather than eccrine, apocrine or follicular gland). This is because SGC can be difficult to diagnose without the 'classical features' mentioned above especially if it is poorly differentiated. Differentiation is split into well, moderately or poorly differentiated; increasing de-differentiation correlates with a higher risk of metastasis and recurrence. The presence of poorly differentiated carcinoma is a high-risk feature irrespective of the amount seen. Thickness of greater than 2 mm upgrades the staging from T1 to T2 and confers a higher risk.(Edge SB, 2010) Lymphovascular invasion and perineural invasion are noted with increased chance of local recurrence and metastasis. Other risk factors include vascular and lymphatic invasion, orbital invasion, involvement of upper and lower eyelids, poor differentiation, multicentric origin, diameter greater than 10 mm, infiltrative growth pattern and pagetoid invasion of the adjacent epithelium.(Rao et al., 1982)

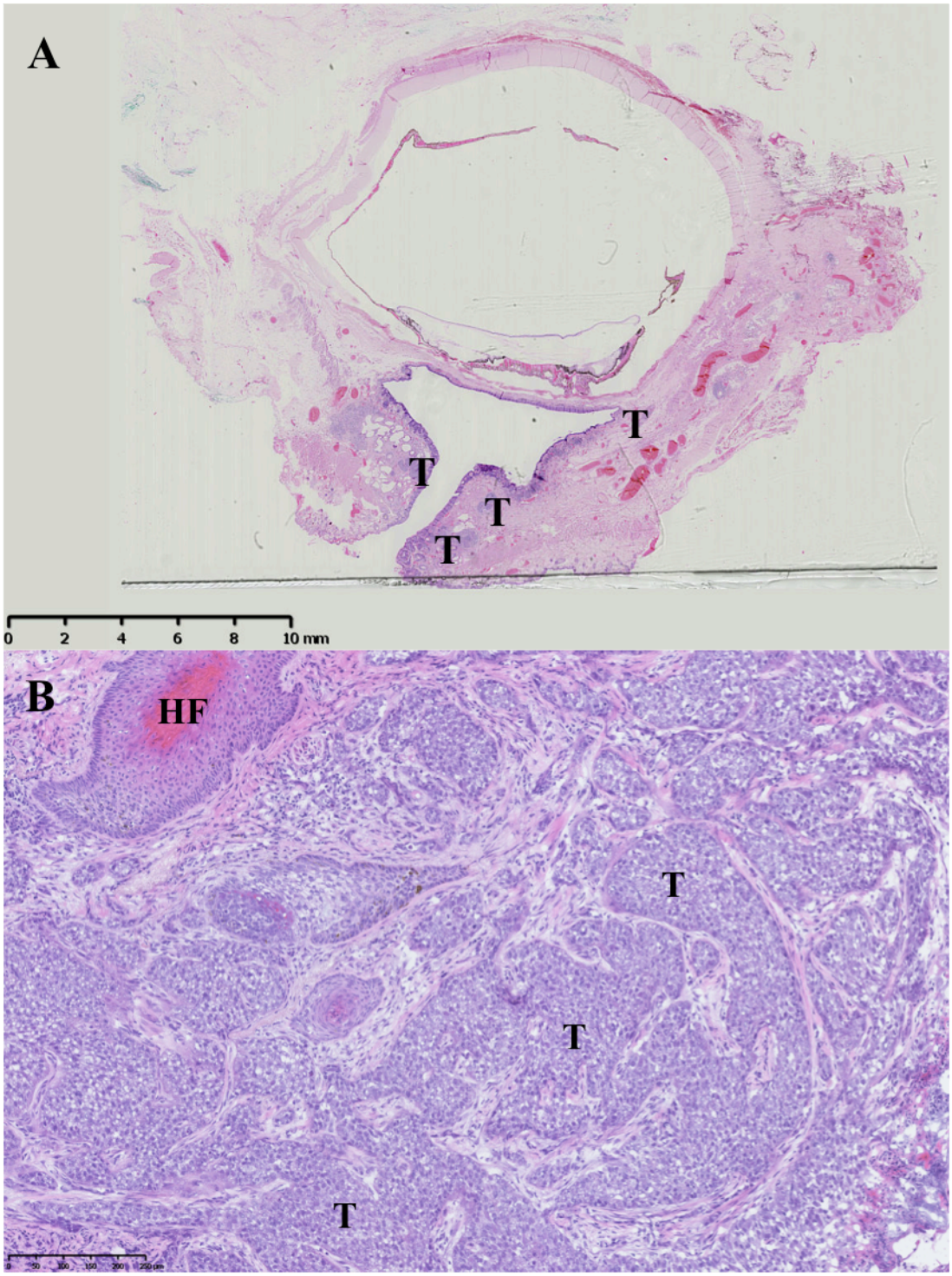


Figure 1.9 Microscopic appearance of SGC. (A) Pagetoid SGC demonstrating the various skip lesions on the conjunctiva of both upper and lower eyelids in this exenterated specimen. (B) SGC highlighting large deranged cells with classic features of intracytoplasmic fat.

1.4.3 Management

Unlike BCC, more intensive staging of the tumour is required to see if it has extensive local spread or metastasis. The principles of treatment are to save life, save vision, save

normal tissue (best cosmetic outcome) whilst saving dignity as the surgery is mutilating. Assessment of the regional lymph nodes using ultrasound, computerised tomography (CT) scanning of the orbit, magnetic resonance imaging (MRI) of the head and thorax and positron emission tomography (PET) scanning are some of the radiological options for staging the disease. If it is localised to the eyelid, then a wide local excision may be all that is required to clear the disease. With extensive disease, removal of the whole globe and orbital content, exenteration, may be required. If the eyelid skin is not involved (i.e. just the conjunctiva and globe) then a skin sparing exenteration is more cosmetically acceptable and probably just as safe.(Shields et al., 2001) Furthermore, this may be combined with sampling of the regional lymph nodes such as a sentinel node biopsy. Local alternative treatment to surgery is radiotherapy although there is controversy to how radiosensitive SGC is.(Harvey and Anderson, 1982, Hendley et al., 1979, Nunery et al., 1983, Matsumoto et al., 1995, Pardo et al., 1989) Newer techniques of radiation delivery such as target splitting non-coplanar RapidArc may be more effective.(Boeva et al., 2014) Distant metastasis requires chemotherapy, although it is not clear which regimen or length of time is best. Cisplatin and doxorubicin have been successful in one patient with lung metastasis.(Koyama et al., 1994) Neoadjuvant chemotherapy with 5-fluorouracil and carboplatin has also been used with success and prevented lymph node dissection. (Murthy et al., 2005, Lisman et al., 1989)

1.4.4 Prognosis and spread

Despite aggressive treatment there is a high recurrence rate and metastasis often occurs via the lymphatics to the cervical lymph nodes.(Buitrago and Joseph, 2008) Mortality rates were once very high at 29%, but there is a significant range depending on which treatment modality is employed and in a can be good as 6% in a centre of excellence.(Doxanas and Green, 1984, Shields et al., 2004, Ni, 1982)

1.5 Genetic basis for cancer

Cancer is a genetic disease. The discovery of the DNA blueprint for life in 1953 led to many theories that mutations transform normal DNA into cancer supporting DNA. These mutations can be induced by either environmental factors (UV light, radiation), local endogenous factors (reactive oxygen or nitrogen species), be inherited in the germline or a combination of all three. The publication of the human genome project in 2001 is one of the greatest milestones in understanding the human blueprint and allows us to classify these mutations involved in carcinogenesis.(Lander et al., 2001, Venter et al., 2001) It is true that in some cases the initiating factor is not a dysregulated gene. For example, hepatitis C initiates hepatocellular cancer, Epstein Barr virus initiates Burkett lymphoma, asbestos initiates mesothelioma. These external, environmental triggers (carcinogens) do effect the genetic makeup and ultimately the progression of the tumour is genetic in nature.

Mutations occur spontaneously at 10^{-8} mutations per base per cell division; with an estimate of 10^{16} cell divisions of our body over a lifetime, then each base within the genome has roughly undergone 10^{8} mutations.(Drake et al., 1998) Clearly, we don't all develop cancer (although 1 in 3 will) thanks to the check and controls that either repair or remove aberrant DNA. With increasing time, mutations accumulate within the cell and the chance of a pathological mutation escaping the normal repair mechanisms increases, which is why age is such a strong risk factor in so many tumours. Interestingly, germline mutation rates differ from somatic cells with the former actually being lower than the latter by almost a half at 10^{-12}.(Orr, 1995)

Not all mutations are pathogenic, the vast majority are synonymous single base change whereby the replacement base makes no alteration in the translated protein (Table 1.6). Mutations within the DNA encompass a spectrum, ranging from a single base to huge chromosomal alterations. There can be confusion with terminology and there are many ways to describe a mutation. Therefore, basic genetic and mutation terminology is summarised in Tables 1.6 and 1.7. For example, a point mutation means something has happened to a single base, but it doesn't tell you about functional outcome. Synonymous and non-synonymous change informs that there is no change in protein, change in protein respectively, however, what is the functional outcome of the non-synonymous change.

The most ideal layout would be a description of the mutation's physical change followed by a functional report. The latter does however require more investigation with either experimentation or the use of predictive algorithms; hence the functional significance of the mutation is not always provided. Conventional views on spontaneous mutagenesis suggests it is a random event with natural selection shaping the evolution of change, though more recently it has been found that the mutational rate varies considerably depending on genome location and is influenced by local genetic and epigenetic features.(Donley and Thayer, 2013)

Terminology	**Description**
Allele (allelomorph)	Alternative form of the same gene located at the same chromosomal position or genetic locus. Being diploid, we have 2 alleles at each genetic locus, normally inherited from each parent
Genotype	Actual sequence of the gene
Phenotype	Physical characteristic of a human or organism
Genetic heterogeneity	Similar or identical phenotypes produced by different genotypes
Pleiotropy	A single gene that causes multiple phenotypes
Polymorphism	Variation in the general population that occurs more or equal to 1%
Transition	Replacement of a purine by another purine or pyrimidine by another pyrimidine
Transversion	Replacement of a purine by a pyrimidine (or vice versa)
Germline mutation	Mutation within the germ cell (sperm/oocyte) making them heritable
Somatic mutation	Acquired mutation occurring in a cell (often referring to the cancer cell) after the first division of the embryo. It cannot be passed onto to its progeny
Driver mutation	A mutation actively involved in tumour growth or progression. Deciphering the early driver mutation is the dream of the cancer researcher
Passenger mutation	A mutation that plays no active role in tumour development. Nevertheless, passenger mutation can become driver mutations during the evolution of the tumour (and vice versa).
Conserved	Same amino acid is coded or tolerable change in amino acid sequence
Null	No active product produced
Gain of function	Hyperactive or new function of a gene or protein product
Loss of function	Reduced or no function of a gene or protein product

Table 1.6 Basic genetic vocabulary.

Genes involved in cancer can be broadly described by their functional phenotype into three categories namely the oncogene, its nemesis, the tumour suppressor gene (TSG) and the stability gene. An oncogene is an inappropriately expressed gene that leads to carcinogenesis. A proto-oncogene is the gene in its prior, controlled normal state. Well

known oncogene products include growth receptors (for example epidermal growth factors, EGF), intracellular signalling molecules (for example tyrosine kinase), cell cycle regulators (for example cyclins) and survival factors (for example apoptosis regulator Bcl-2, BCL2). Their products can be any type of cellular protein. A collection of mutations in growth promoting genes (oncogenes) is thought to happen early in cancer. A good example is avian myelocytoviral oncogene (*MYC*) proto-oncogene that encodes a DNA-binding factor, located on 8q24.21, that can activate transcription.(Dang et al., 1999) Finding these early oncogenes (the early driver mutations) specific to the cancer variety will assist in early intervention with a view to cure. Timing is everything, as the more the tumour proliferates, the more mutations it collects causing it to 'branch' away from the tissue of origin (germline DNA) so much so that modification of the early 'trunk' mutations may no longer have a significant impact on the course of disease. Tumours contain millions of cells and the studying the origin and evolution of these cells is termed cancer phylogenetics.

TSGs provide the appropriate stop mechanism to normally proliferating cells. Their normal function is to control the cell cycle and apoptosis. In contrast to oncogene activation, it is 'loss' of the tumour suppressor gene that allows uncontrolled growth. Presence of these 'suppressor products' was demonstrated in an elegant experiment where malignant cells that were fused with normal cells started to lose their tumorigenic behaviour.(Harris et al., 1969) Over time, loss of certain part of the chromosome in the normal cell allowed the return of their tumorigenic behaviour.

The paradigm tumour suppressor gene is the retinoblastoma 1 (*RB1*) gene found on chromosome (Chr) 13q14.1-q14.2 and causes the hereditary childhood cancer, retinoblastoma (RB). The RB1 protein binds to the E2F family of transcription factors and represses the transcription of genes needed for the S phase of the cell cycle, thus is a repressor of the cell cycle. (Nevins, 2001) Furthermore, RB is a paradigm for the Knudson's two hit hypothesis whereby loss of both *RB1* genes are required for the tumour to develop.(Cavenee et al., 1983) RB can be broadly split into germline (autosomal dominant) and sporadic RB cases, although only 10% of children have a positive family history: the majority of germline cases are new mutations. Overall 40% of cases carry a germline mutation of the *RB1* gene with the rest been sporadic. Sporadic mutations only occur in one primitive retinal cell (retinoblast) making them unilateral in nature (only

affecting one eye) and have no risk of passing it onto their offspring. A germline, heterozygous mutation is present in all cells of the body and so the chance of each eye containing thousands of retinal cells getting a second hit is high and so tend to present with multifocal or bilateral (affecting both eyes) tumours. Nevertheless, unilateral cases do not exclude the presence of a germline mutation and 15% of unilateral cases are germline in nature, thus genetic screening is done on all patients to determine risk of inheritance and prognosis. All germline tumour patients are at risk of secondary tumours especially osteosarcoma, and soft tissue sarcoma in up to 19% of cases.(Derkinderen et al., 1987) External beam radiation used to be a mainstay treatment for RB, but fortunately alternative therapies are available, subsequently reducing the risk of future tumours. The risk to offspring of germline carrier *RB1* mutations of developing RB is 45% as the mutation is non-penetrant in 10% of carriers.

Other common tumour suppressor genes have been discovered in a similar fashion to RB, as a result of their specific involvement in familial cancer syndromes: *BRCA1/2* in breast/ovarian cancer and *P53* in Li Fraumeni familial cancer syndrome. P53 is also inactivated by somatic mutations and is often found in many types of cancer. *TP53,* via *p53* expression, controls DNA synthesis, apoptosis, angiogenesis and senescence in normal cells. Loss of p53 expression allows the aforementioned processes to go unimpeded and ultimately promote tumour development.

Aberrant stability genes (caretaker genes) don't promote growth directly, instead they lead to either genomic instability or a chromosomal rearrangement that allows an increase in mutation rate and uncontrolled replication. Examples of genes involved in stability include nucleotide or base excision repair, mismatch repair, telomere metabolism and non homologous end joining genes. Moreover, replication timing in somatic cells influence the rate of mutagenesis and may contribute to the development or progression of cancer. Genome instability can lead to a higher or lower than normal number of chromosomes or aneuploidy. Another phenomenon that causes genetic change is chromothripsis. This is where the genome undergoes a single catastrophic event whereby one or several chromosomes shatter into hundreds of fragments and the DNA repair machinery recombines these fragments in an erroneous fashion. Massive gene amplifications or deletions can occur to multiple oncogenes or TSGs in a single event. The exact definition of chromothripsis is not in agreement, but as a minimum includes 10 copy number

variations (CNV) or loss of heterozygosity (LOH) transitions on a single chromosome arm. CNV in cancer is termed copy number aberration (CNA) and plays an important role in altering genetic function. Detection of this is traditionally done using karyotyping where multiple copies or loss of a chromosome is detected using photographs of dividing cells arrested in metaphase and stained with Giemsa. Evolution of the technique resulted in multicolour fluorescent in-situ hybridisation or spectral karyotyping where each chromosome is given its own colour which makes it easier to pick up chromosomal rearrangements especially if they are balanced. Comparative genomic hybridisation allows the patient normal DNA (say blood) and the tumour DNA to compete with binding to normal chromosomes, each with a different colour, say red/green respectively. Unaltered regions are demonstrated with equal binding (thus colour) for that region. If the region is amplified in the tumour, more green colour would be seen. Newer bioinformatic programs have allowed the mining of whole exome sequencing (WES) to detect CNA and LOH. This allows the appropriate interrogation of WES data: a variant detected may only be on one allele of the chromosome, but if the opposing allele is lost or there are multiple copies of that allele then it will have more of an effect. We used three two techniques to interrogate our WES including ASCAT, DNA copy and Varscan 2 (see Chapter 2).

Descriptive terminology	**Variant or mutation**	**Alternative terminology**
Synonymous single nucleotide variant (snv)	Single base change with same protein	Synonymous substitution
Non-synonymous snv	Single base change with different protein/no protein	Non-synonymous substitution Nonsense if stop codon Null mutation if no product Mis-sense mutation if one amino acid replaced by another Point mutation
Aneuploidy	Change in number of whole chromosomes	
Sequence inversion, duplication, or translocation	Rearrangement of chromosomal material	
Deletion	Deletion of any number of bases	
Insertion	Insertion of any number of bases	
Frameshift	Insertion/deletion that alters the triplet reading frame	
Amplification	Increased number of copies of gene coding sequence	
Copy number variation	Increase in DNA sequence (usually large)	DNA copy number (more copies of a gene within the DNA)
Loss of heterozygosity	Loss of one copy of an allele	
Mosacisim	Cells within the same individual contain different genotypes	Somatic mosacisim: non-gamete cells contain different genotypes Germline mosacisim: gamete cells contan different genotypes

Table 1.7 Basic terminologies for describing DNA mutations change in code.

Finding the relevant early driver mutations is the desire of cancer geneticists and it may then be possible to modify the genes in the future to halt or stem the disease. These critical early mutations are significant enough to provide a growth advantage, induce enough instability in a relatively normal cell that have defence mechanisms and thus there are probably only a few ways it which the mutation can induce uncontrolled proliferation at the initiation stage. Other genes are activated at later stages as the cancer diverges from normal, although they could have been present at an early stage (termed passenger mutations), but had no biological value to the proliferating tumour at that time. These passenger mutations, for reasons just mentioned, should not be totally disregarded. Identification of critical mutations is often done by comparing the DNA of the tumour with the patients germline DNA, say from a blood sample (assuming that it is not a haematopoietic cancer; you could use saliva in such cases). Any mutations identified only in the tumour DNA are deemed relevant, however, they still need to be grouped into harmful or harmless changes. These genetic mechanisms of tumour progression fits well with the clonal expansion theory whereby mutant tumour cells have a growth advantage, expand and become the dominant cell population.

Changes do not just occur at the somatic or germline genome level, but also happen at an epigenetic level too. Epigenetics is about the way the DNA is packaged into chromatin. This involves methylation of DNA, histone modification, nucleosome positioning and spatial arrangement of the chromatin. This spatial arrangement allows unpackaging of the DNA at certain areas so that only a specific set of genes in the genome are expressed. The more specialized the cell, the more epigenetic controls are required. These epigenetic controls are passed on to daughter cells and are heritable to allow the developing embryo to form, however, whether alterations in these controls can be passed on from parent to child is unclear. Cancer gains reprogrammed epigenetic controls that confer an advantage. DNA methylation only occurs (in humans) at cytosine bases located 5' to a guanosine in a CpG dinucleotide, a so called CpG island.(Bird, 2002) These islands are both hypo- and hypermethylated in cancer. As almost half of all genes have a CpG island associated with their promotor, the tumour takes advantage of this. Promotor methylation can silent certain genes and silencing of TSGs occurs in tumours. Interestingly, some genes are only silence by methylation, others only by genomic mutation and others still by both mechanisms. For example, MutS Homolog 2 (*MSH2)* is never silenced epigenetically whereas MutL Homolog 1 (*MLH1)* is often epigenetically silenced.(Jones

and Baylin, 2002, Baylin and Jones, 2011) In contrast, hypomethylation may induce chromosomal instability. A newer concept in epigenetics are topographically associated domains (TADs) whereby the human genome is organised into regions of megabases with distinct boundaries.(Lupianez et al., 2015) Enhancers and promotors act within a TAD and the boundaries act as a shield to those genes outside of this region. Re-arrangements of these boundaries can cause enhances to act upon genes outside of their original TAD. Loss of these boundaries occur with so called 'balanced' CNV or non-coding DNA CNV and go on to explain why these changes are tumourigenic.

The ability of tumours to maintain a proliferative state is thought to be related to some cancer cells having the ability of self renewal and behaving like a pluripotent stem cell. The cancer stem cell theory goes on to say that only a few cancer cells have this ability and are there from the very beginning, thus are the initiators of tumorigenesis, but also maintains cell division as the tumour progresses. The theory is more robust for haematologically malignancies where first evidence came the transplantation of acute myeloid leukaemia in mice using a few cells only.(Lapidot et al., 1994) This built on evidence of tumour transplantation demonstrated as early as 1937.(Southam and Brunschwig, 1961, Hewitt, 1958) Solid tumours have also demonstrated the existence of cancer stem cells with it first identified in breast cancer.(Al-Hajj et al., 2003) Uncontrolled self renewal is an attractive proposal, however, how they arise in the first place and what mechanisms allow their pluripotency are just starting to be unravelled. The stem cell theory and clonal expansion theory do not have to be mutually exclusive and it is more likely that both mechanisms are acting within cancer.

MicoRNA (miRNA) are short non-coding RNA 22 nucelotide sequences that function as post-transcriptional regulation of gene expression in both the physiological and pathological state. MiRNA are described with the prefix as the organism, hsa- being homo sapiens. Naming of the gene uses a small r, for example hsa-mir-100, whereas the mature 22 nucleotide sequence has a large R, namely hsa-miR-100. Sequences can be derived from both the 3' and 5' hairpin precursor with increasing evidence to suggest that both are biologically active. They are described with the suffix -3p or -5p. Identical miRNA sequences can arise from different genomic loci and a numerical suffix -1, -2 is added to distinguish between them. Very similar sequences which are only differ by a few bases have a small alphabetic letter suffix to the main numerical identifier, for example hsa-

100a and hsa-100b. The public repository miRBase is useful resource to correctly describe and discover their genomic location.(Kozomara and Griffiths-Jones, 2014) MiRNA and epigenetic factors are regulators of gene expression. MiRNA undergo epigenetic modification under that same constraints as protein coding genes, however, miRNA can directly target epigenetic machinery such as DNA methyltransferases or histone deacetylases. This epigenetic-miRNA regulatory circuit can be disturbed in tumours and adds another layer of complexity when trying to dissect carcinogenesis.(Sato et al., 2011) MiRNA binds to messenger RNA and depending on their degree of base pair homology (Watsonian match), the mRNA undergoes cleavage or translational repression. This causes silencing of mRNA expression, however, it can also increase the expression of mRNA. Moreover, they have been shown to target DNA directly. Genomic regions associated with cancer that undergo LOH, amplification and deletion also contain miRNA plus the alteration of the miRNome has been shown in several cancers.(Calin 2004Proc , calin 2006Nat rev) In a similar fashion to gene, oncogenic miRNA are labelled oncomiR and others, such as hsa-miR-199a, have been found to be tumour suppressive.(Cho, 2007, Cheng et al., 2012)

1.6 Basal cell carcinoma genetics

The genetics of BCC has been alluded to by the analysis of basal cell naevus syndrome (BCNS, aka Gorlin syndrome). This is a condition where patients develop multiple BCCs, medulloblastoma and craniofacial defects. It is an autosomal dominant, heterozygous germline deletion of the gene Patched 1 receptor (*PTCH1*) located on Chr9q.22.32 (Table 1.8).(Lorenz and Fuhrmann, 1978) The location of the gene was initially identified using linkage analysis and narrowed down to Ch9q22.3-q31.(Farndon et al., 1992) The human homologue of Drosophila PTCH was then identified as the likely candidate for BCNS and sporadic BCC.(Gailani et al., 1996, Johnson et al., 1996) Moreover, Knudson's two hit hypothesis was suggested to apply to the all components of BCNS, however, hemizygosity is enough for the craniofacial malformations, but homozygosity is required for cancer development.(Levanat et al., 1996) *PTCH1* is essentially a tumour suppressor gene as it role is to maintain the Hh pathway in the off state when no ligand is present. Loss of this repression will lead to uncontrolled and continuous activation of the pathway. The heterozygous mutation is sufficient to induce the skeletal defects during embryogenesis, however, the genesis of BCC requires a second hit (the Knudson's two-

hit hypothesis). An environmental insult is thought to cause the mutation in the second allele (such as UV light damage), however, the majority will be compound heterozygous (same gene, but not exactly the same location). This implicates that dysregulation in the hedgehog (Hh) signalling pathway is an early event in the development of BCC. Inherited BCC only account for 1-2% of the BCC workload and the vast majority arise sporadically. Sporadic BCC have shown heterozygous *PTCH 1*and *PTCH2* mutations, however, it is not universal.(Jayaraman et al., 2014) Nonetheless, other parts of the Hh signalling pathway have been mutationally affected such as smoothened (SMO).(Otsuka et al., 2015, Pricl et al., 2015)

Disease	**Defective gene**	**Cytogenetic locality**
Albinism	Tyrosinase (*TYR*)	11q14.3
	Oculocutaneous Albinism II (*OCA2*)	15q13.1
	Tyrosinase-Related Protein 1 (*TYRP1*)	9p23
	Solute Carrier Family 45, Member 2 (*SLC45A2*)	5p13.2
	Hermansky-Pudlak Syndrome 1 (*HPS1*)	10q24.2
	Lysosomal Trafficking Regulator (*LYST*)	1q42.3
BCNS	Patched receptor 1 (*PTCH1)*	9q22.32
	Patched receptor 1 (*PTCH2*)	1p34.1
Bazex's syndrome	Unknown	Xq24-q27
Rombo syndrome	Unknown: Possibly a DNA repair or cell cycle regulation gene	Unknown
Xeroderma pigmentosa	Excision Repair Cross-Complementation Group 4 (*ERCC4*)	16p13.12

Table 1.8 Syndromes associated with BCC highlighting genetic heterogeneity

1.7 Sebaceous gland carcinoma genetics

The genetics of sebaceous gland carcinoma is essentially unknown. Muir-Torre syndrome is where alterations in the DNA mismatch repair genes *MSH2* located on Chromosome 2 (2p21) or *MLH1* located on Chromosome 3 (3p22.5) lead to cutaneous and visceral tumours.(Schwartz and Torre, 1995) These cutaneous tumours include sebaceous adenoma, keratocanthoma, BCC, and visceral include colorectal adenocarcinoma and genitourinary malignancies. It resembles a hereditary non-polyposis colorectal cancer syndrome (HNPCC), and is autosomal dominant in nature. HNPCC involves changes to DNA microsatellites. These microsatellites are short repeats of DNA that are often read incorrectly by DNA polymerase. Proof reading mechanisms (also known as mismatch repair mechanisms) have to compare the newly synthesized (sense) strand to the template

strand to pick these incorrect reads made by the DNA polymerase and repair them. Failure of these mismatch repair (MMR) gene products to pick this up, leads to HNPCC. There are six MMR genes and listed in Table 1.9. *MSH2* is the most frequent change seen in HNPCC. Although commonly associated with sebaceous gland carcinoma, its true association is actually rare (sebaceous adenoma is the true association): a case series of 60 revealed only one case of Muir Torre syndrome, highlighting that this (SGC) is unlikely to be a syndrome condition.(Shields et al., 2004, Svajdler et al., 2015) Moreover, sebaceous adenoma does not seem to convert into sebaceous carcinoma (i.e. it is not a premalignant step to the development of SGC).(Lazar et al., 2007) Nevertheless, in up to 40% of cases in MTS, the sebaceous adenoma predates the internal malignancy and so it is still recommended to screen for mismatch repair gene products. Furthermore, immunostaining of the sebaceous adenoma for MLH1 and MSH2 proteins is a valid tool to help diagnose MTS.(Jagan, 2014)

Periocular SGC has not been readily investigated. Kim et al carried out some immunostaining comparing metastatic with non-metastatic SGC to find that the metastatic group demonstrated higher expression levels of sonic hedgehog (SHH), wnt and ATP-Binding Cassette, Sub-Family G, Member 2 (ABCG2).(Kim et al., 2013d) A retrospective study looked at 16 cases and using immunostaining identified increased expression of b-catenin, cellular retinoic acid binding protein 2 (CRABP2) and fatty acid-binding protein 5 (FABP5) compared to controls.(Tsai et al., 2015) Over expression of ZEB2 by immunostaining was associated with lymph node metastasis and poorer prognosis in a 65 case study.(Bhardwaj et al., 2015) Fluorescence in situ hybridization highlighted extra copies of two epidermal growth factor (EGF) receptors, namely epidermal growth factor receptor (*EGFR*) and Erb-B2 Receptor Tyrosine Kinase 2 (*ERBB2*) along with fibroblast growth factor receptor 1 (*FGFR1*) 3-6 of 49 SGC cases.(Lee et al., 2015)

Gene	**Cytogenetic locality**
MutS Homolog 6 (*MSH6)*	2p16.3
MutS Homolog 2 (*MSH2*)	2p21
PMS1 Postmeiotic Segregation Increased 1 (S. Cerevisiae) (*PMS1*)	2q32.2
MutL Homolog 1 (*MLH1*)	3p22.5
PMS2 Postmeiotic Segregation Increased 2 (S. Cerevisiae) (*PMS2*)	7p22.1
MutL Homolog 3 (*MLH3*)	14q24.3

Table 1.9 Mismatch repair genes and their cytogenetic location

1.8 Cancer signalling pathways

Genes are functional units of DNA. This functional code produces a transcript (primary RNA) in the nucleus, becomes modified by splicing the introns out to produce a messenger RNA (mRNA) that moves into the cytoplasm. The primary structure of RNA is similar to DNA except that it is single stranded with ribose instead of deoxyribose and thymine (T) is replaced by uracil (U). MRNA is composed of trinucleotide repeats (codons) which produce 64 combinations; of these 3 encode stop codons (UGA, UAG, and UAA) and the other 61 encode 20 amino acids. Interestingly, the genetic code is universal across all species - the same codons specify the same amino acids in all of life. Aminoacyl-transfer RNA (tRNA) reads the mRNA cipher within the ribosome to align the correct amino acids to form the polypeptide chain. The polypeptide undergoes many post-translational modification steps including folding, chemical modification and transportation before forming a protein. Moreover, one gene can code for more than one protein as a result of splicing out different parts of the primary RNA. Furthermore, genes are switched on and off depending on the state of the cell. Some genes code for pathways that are essential for embryogenesis, but are switched off in most adult cells. Hh, EGF and wingless-type MMTV (wnt) are examples of such pathways that are highly active during embryogenesis, but most of the time lay dormant in adult cells. These embrogenic pathways promote growth and the laying down of new tissues, hence in cancer these proliferative pathways are reactivated.

1.8.1 Hedgehog signalling pathway

Hh is a critical embryonic pathway whose role is essential for temporal and positional development of tissue such as the appendages. It is switched off in the majority of adult cells bar transient expression in hair follicles, reproductive organs and parts of the brain along with a role in stem cell maintenance.(Petrova and Joyner, 2014, Sahin et al., 2014, Morales et al., 2009) Three Hh protein homolog's bind to PTCH1, a cell membrane receptor, to initiate the canonical Hedgehog (Hh) pathway. SHH, indian hedgehog (IHH) and desert hedghog (DHH) are the three homologs. SHH is involved in polarising limb buds and notocord organisation; IHH controls cartilage and bone development; DHH is required for germ cell development and nerve sheath construction.(Briscoe and Therond, 2013) In the absence of the ligand, PTCH1 represses the activation of SMO, a receptor

on the membrane of intracellular endosomes, effectively keeping it in the off position. Binding of the ligand removes PTCH1 repression allowing the translocation of SMO from the endosome to the primary cilium where it activates the Glioma-associated oncogene (Gli) zinc finger transcription factors 1,2 and 3 (Figure 1.10). Gli then enter the nucleus to activate or repress target genes. Gli1 is mainly described as an activator, Gli3 a suppressor, and Gli2 as having dual function.(Briscoe and Therond, 2013)

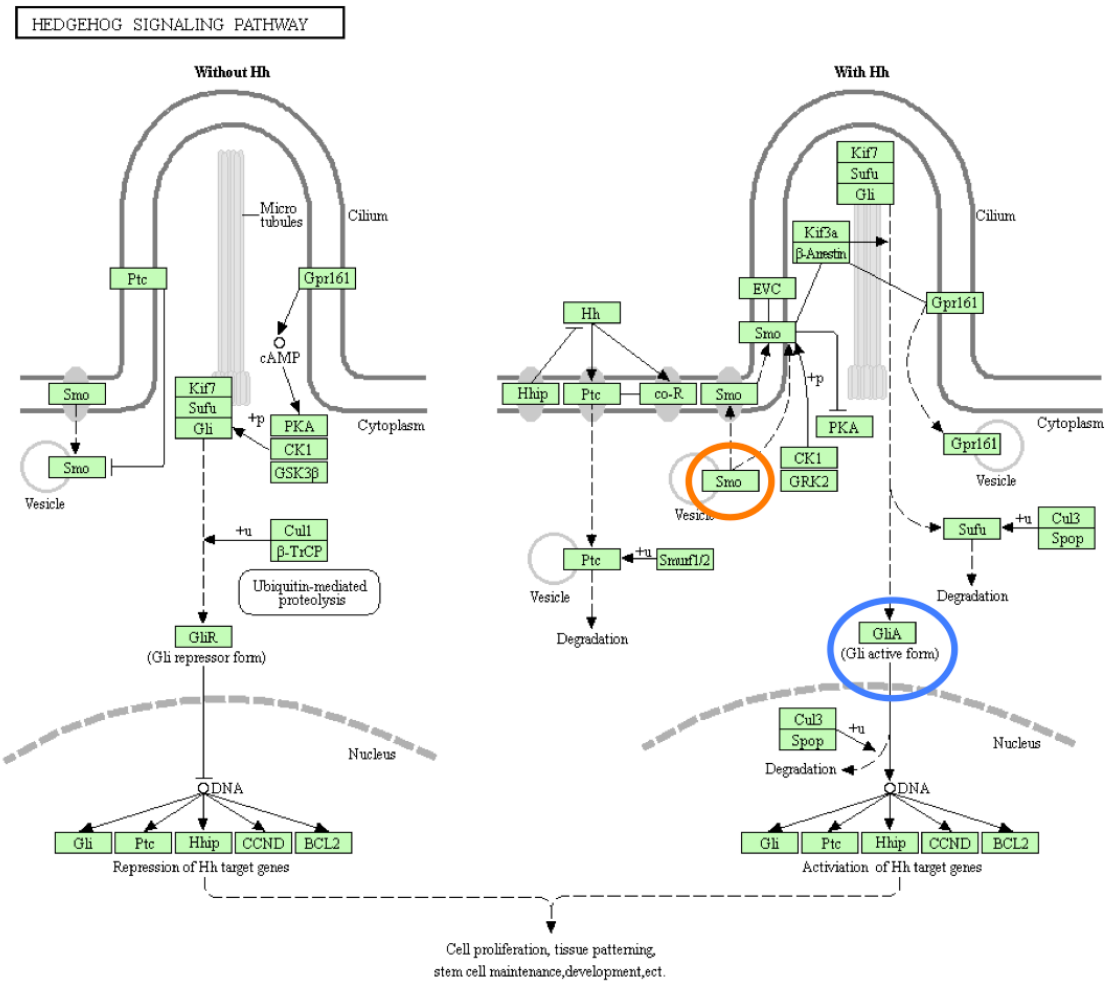


Figure 1.10 Canonical Hedgehog signalling pathway. Orange circle denotes available therapeutic agents, namely SMO inhibitors. Blue circle denotes newer therapies being developed for more downstream inhibition of the Hh pathway (Adapted from KEGG pathway)

SMO promotes the dissociation of the Supressor-of-fused (SUFU)-Gli complex. Three kinases regulate Gli processing within the primary cilia: protein kinase A (PKA), casein kinase 1α (CK1) and glycogen synthase kinase 3β (GSK3B). Kinesin Family Member 7 (KIF7) partly controls movement around the primary cilia. All Gli have C-terminal activation domains, but only Gli2 and 3 have N-terminal repressor domains. The activating C-terminal domain undergoes phosphorylation and partial proteolytic cleavage

in the ‘off state’ allowing the N-terminal and DNA blinding sites to translocate to the nucleus to repress transcription. ‘On state’ (binding of the Hh ligand) inhibits phosphorylation and proteolytic cleavage, allowing C-terminal activator domains to become functional.(Briscoe and Therond, 2013)

1.8.2 Hedgehog activation in cancer

Pathological activation of Hh signalling occurs in a variety of cancers, including inherited BCC, prostate cancer, upper gastrointestinal tumours, a subset of small cell lung carcinoma and pancreatic ductal adenocarcinoma; the role being different in each: initiation of tumour, excess ligand being essential for tumour growth, cancer cell survival, maintenance of malignant phenotype, and both early and late regulation of tumorigenesis respectively.(Berman et al., 2003, Watkins et al., 2003, Thayer et al., 2003, Dahmane et al., 1997, Sanchez et al., 2005) BCNS highlights the cancer initiator role of the Hh pathway, presumably by its role in stem cell maintenance of the hair follicle and Purkinje cells where BCCs and medulloblastomas are thought to arise respectively. Gli-1 over expression is seen in both inherited and sporadic BCCs.(Dahmane et al., 1997) The diversity of Hh involvement reflects it varied function, but also means there is no simple solution or modification of the Hh pathway that would help all tumours as reflected by discouraging results seen by cancer trials with Hh pathway inhibitors.(Gonnissen et al., 2015) Where no Hh pathway mutation is found, the role of Hh in such cases could be to promote the Warburg glycolytic metabolism, to help support tumour development or promote tumour microenvironment possibly via a stromal paracrine signalling effects.(Teperino et al., 2012, Yauch et al., 2008)

Much excitement has surrounded the development of SMO inhibitors to treat Hh pathway associated cancers. Vismodegib was shown to successfully in patients with BCNS: it reduced the burden of skin BCCs clinically and showed a 90% reduction in Gli-1 mRNA levels at one month post-treatment.(Sekulic et al., 2012) Unfortunately, 54% discontinued drug usage due to side effects including loss of taste, hair, weight and muscle cramps. Good initial responses are seen with the drug, but secondary resistance has been noted.(Pricl et al., 2014, Pricl et al., 2015) Vismodegib has also been used in metastatic pancreatic adenocarcinoma with less success, although at a molecular level it showed reduced levels of Gli-1 and PTCH1, the pool of cancer stem cells remained and at a

clinical level, survival was inferior to a more traditional chemotherapy drug Gemcitabine.(Kim et al., 2014)

1.8.3 Epidermal Growth factor (EGF) pathway

EGF pathway is an essential pathway for cell growth, differentiation and proliferation both in the developing embryo and adult. A family of tyrosine kinase receptor mediates its action, namely EGFR, ERBB2, erb-B2 receptor tyrosine kinase 3 (ERBB3) and erb-B2 receptor tyrosine kinase 4 (ERBB4) or more commonly known as human epidermal receptor (HER)-1, HER2, HER-3 and HER-4. A number of different ligands can bind to each receptor that resides on the plasma membrane. Dimerisation of the receptor occurs to trigger a cascade of events depending on the ligand/receptor involved. Events comprise of protein synthesis, glycolysis and via nuclear localization, gene regulation, kinase function and protein interactions. EGF receptors can even translocate to the mitochondrion to activate cytochrome c oxidase subunit II; a critical pathway of oxidative phosphorylation.(Boerner et al., 2004) The commonest ligands are EGF, heparin-binding EGF-like growth factor and transforming growth factor-α. Activated EGFR can signal via a kinase-dependent or kinase-independent route. The former signals phospholipase-Cγ/protein kinase C (PLC-γ/PKC); rat sarcoma viral oncogene homolog/rapidly accelerated fibrosarcoma kinase/mitogen activated protein kinase kinase/mitogen activated protein kinase (Ras/Raf/MEK/MAPK); phosphatidylinositol 3-kinase/ RAC-alpha serine-threonine-protein kinase/ Mechanistic Target Of Rapamycin (PI3K/AKT/mTOR); Janus Kinase 2/ Signal Transducer And Activator Of Transcription 3 (JAK2/STAT3) and STAT3 (independent of JAK2). The latter signals MAPK, AKT pathway, Solute Carrier Family 5 (Sodium/Glucose Cotransporter), Member 1 (SLC5A1) and BCL2 Binding Component 3 (BBC3) usually by direct interaction with proteins.

1.8.4 EGF pathway in cancer

Several cancers have demonstrated increased activity of EGF pathways via mutation of the EGFR, upregulation of the EGFR or via a receptor independent route. EGFR mediated carcinogenesis is summarised in Table 1.10. Mutations are seen in the tyrosine kinase portion of the EGFR in a subtype of lung cancer, non-small cell lung cancer

(NSCLC).(Bunn and Franklin, 2002) Being on the plasma membrane, targeted therapy using antibodies and tyrosine kinase small molecule inhibitor. Erlotinib and Gefitinib are the first-generation tyrosine kinase inhibitors. Erlotinib versus carboplatin-gemcitabine chemotherapy in NSCLC demonstrated progression free survival (PFS) rates of 13.1 months versus 4.6 months respectively.(Zhou et al., 2011) Gefitinib versus carboplatin-paclitaxel chemotherapy demonstrated PFS rates at 12 months of 25% versus 7% respectively.(Mok et al., 2009) Subsequently, similar EGFR changes were noted in certain types of breast, colorectal cancer and pancreatic cancer, where the drugs have been used.

Pathways do not act in isolation, and cross talk between pathways is likely to constantly be happening. This adds a layer of complexity when trying to dissect out important driver pathways. An example of crosstalk in BCC was demonstrated by an over expressing Gli1 (Hh signalling) cell line model of nodular BCC, that repressed MAPK1 (EGF signalling) which kept the cell nodular like colonies in an epithelial phenotype preventing migration.(Neill et al., 2008)

Receptor	Pathway and action
EGFR	Ras/Raf/MEK/MAPK pathway: cell proliferation
	FAK pathway: tumour migration
	STAT3 pathway: gene transcription
	PI3K pathway: cell survival
ERBB2	HER1/HER2 - MAPK pathway: cell proliferation
	HER2/HER3- PI3K pathway: aggressive phenotype
	HER2/HER3-may confer drug resistance
	Note: there is no known ligand that binds to HER-2, nevertheless, HER-2 is the preferred dimerisation partner with the other receptors
ERBB3	HER2/HER3- PI3K signalling pathway: aggressive phenotype
	HER2/HER3-may confer drug resistance
	Note: HER-3 can bind multiple ligands but cannot carry out kinase activity, requiring dimerisation with another ligand, commonly HER2, to function
ERBB4	Tissue differentiation
	Binds multiple ligands

Table 1.10 Main actions of EGF receptors in carcinogenesis.

1.9 Stromal milieu in cancer

Cancer behaviour is not restricted to the tumour borders delineated on microscopy; the surrounding apparent normal tissue, termed stroma, has been shown to play a key role in local spread.(Davidson et al., 2014) During embryogenesis, communication between epithelial tissue and surrounding stroma occurs by Hh signalling in order to promote growth and differentiation of the stroma: urogenital epithelium utilises Hh signalling to promote surrounding tissue to develop into the prostate gland.(Podlasek et al., 1999) Reactivation of this paracrine phenomenon of Hh signalling has been shown in prostate cancer stroma and the tumour regulates the proliferation of adjacent non-malignant epithelium.(Wilkinson et al., 2013) Studies in ovarian cancer have highlighted the importance of the changes in the stroma that confer a poorer prognosis and the promotion of local invasion in morphoeic BCC has been demonstrated by modulating the stromal milieu.(Marsh et al., 2008, Davidson et al., 2014, Kataoka et al., 2012)

Preliminary in vitro and in vivo work in the lab highlighted that local invasion seen in the morphoeic subtype was promoted by the integrin beta 6 (ITGB6) through activation of transforming growth factor-beta 1 (TGFB1) that modulates the surrounding stroma.(Marsh et al., 2008) Furthermore, TGFB1 may be the culprit of the myofibroblastic phenotype that is seen surrounding the morphoeic tumours.

1.10 Hypothesis and aims of study

Periocular malignancy represents an increasing burden and currently requires disfiguring surgery in attempt to cure patients. Basal cell carcinoma is the commonest cancer worldwide and morphoeic BCC is an aggressive subtype requiring more intensive surgical excision. Sebaceous gland carcinoma is a rare, but life threatening condition that often requires blinding surgery to prevent mortality. The genetic and molecular characteristics of these tumours have not been studied in detail. For BCC, exome sequencing of 10 BCC elucidated some candidate genes, however, the subtypes were not mentioned which means the majority or all were nodular as these are the most common. The genetics of the aggressive subtype has not been looked at. As mentioned, pilot studies in the Philpott lab on human cell lines have suggested a difference in the Hh expression levels account for the varied phenotype or behaviour between nodular and morphoeic

subtypes. Furthermore, modification of the surrounding apparent normal stromal tissue to promote local spread. Immunostaining has been the mainstay of SGC investigation and recent studies have excess SHH ligand implicating the Hh signalling pathway in SGC carcinogenesis. A two-stage approach was taken to elucidate tumorigenesis. Firstly, a discovery approach identifying the mutational burden along with driver mutations was undertaken. Secondly, a functional analysis of the transcriptome and protein was performed to implicate pathways, including Hh, involved and relate them to the mutational burden.

Hypothesis: The interplay between Hedgehog and other pathway signalling in BCC and SGC, along with differences in tumour and stromal expression underlies the pathogenesis of different histological subtypes and their degree of invasiveness.

Two main approaches were taken to study this:
i) Mutational burden analysis to discover new driver mutations
ii) Transcriptome and protein investigation to identify known and novel pathways and relate this to the mutational discovery

In order to answer this hypothesis, the following aims were formulated:

1) Obtain ethical approval across multi-site to recruit as many patients as possible over the 3 years
2) Create a multi-specialty collaboration between ophthalmology, dermatology, pathology, molecular genetics, molecular biology, bioinformatics
3) Obtain funding for the project
4) Devise protocol for nucleic acid extraction from patient's blood and tumour as well as transfer to the lab
5) Perfect laser capture microdissection to obtain pure tumour
6) Ensure correct bioinformatics pipeline for next generation sequencing analysis
7) Mine data to identify driver mutation, transcriptome changes in relation to known and unknown pathways
8) Carry out verification and validation of data
9) Present and publish data.
10) Use conclusion to devise further work

CHAPTER TWO

METHODS AND MATERIALS

2.1 Human tissue

2.1.1 Ethical approval for the use of human tissue

Prospective approval

The concept, design, and ethical approval was all performed by myself in the preceding two years of the PhD. A new collaboration was forged between Moorfields Eye Hospital (MEH), Institute for Ophthalmology (IoO), Blizard Institute and Barts cancer Institute (BCI). Expansion of ethical approval and collaboration with Guys and St Thomas' NHS Trust (GSTFT) occurred during the PhD.

Ethical approval was granted to acquire prospectively tissue samples of human nodBCC, mBCC and SGC along with blood samples over 3 years from patients attending MEH. MEH Biobank research ethic committee (REC), reference 10/H0106/57-2012ETR28, established the approval. Following informed consent (Figure 2.1) from the patient, a small part of surplus tumour was taken for research during the routine scheduled eyelid cancer surgery, the margins were not compromised and the remaining tissue sent for routine histopathological examination at the IoO. The patient's normal treatment was not affected in any way. Several Moorfield's patients went onto to have Mohs micrographic excision at St Thomas' Hospital (acting as a tissue collection centre) and further surplus tissue was collected if required. Permission for St Thomas' hospital to become a tissue collection centre was gained from GSTFT research and development team and given the local number RJ113/N155.

A proportionate REC amendment was sort to convert St Thomas hospital from a tissue collection centre for MEH patients only into a full recruitment centre for the aforementioned tumours. National Research Ethics Service (NRES) Committee North West - Greater Manchester South conferred a favourable opinion and coded the study as follows 14/NW/1080. Thus, the project recruited in similar fashion patients attending the

Dermatological Surgery and Laser Unit at St Thomas' Hospital for Mohs micrographic surgery directly without a prior visit to Moorfields Eye Hospital.

KING'S HEALTH PARTNERS

Guy's and St Thomas' NHS
NHS Foundation Trust

Queen Mary University of London

Moorfields Eye Hospital NHS
NHS Foundation Trust

St John's Institute of Dermatology
Guy's Hospital
Great Maze Pond
London SE1 9RT

Telephone: 020 718 86273 / 86277 / 89142

Fax 020 718 86113

Email: stru@gstt.nhs.uk

CONSENT FORM

SUBJECT INFORMATION SHEET and INFORMED CONSENT FORM
Title of Study: Identification of genetic factors involved in morphoeic basal cell (mBCC) and sebaceous gland carcinoma (SGC) of human eyelid tumours with a view to identifying potential treatment targets.
Version and Date of Subject Information Sheet/Informed Consent: 1/6/14 Version 2.2
Ethics Committee Code: **Ethics Committee Code**: **14/NW/1080**

Name of Researcher: Dr John Bladen **Please initial box**

1. I confirm that I have read and understand the information sheet dated 1/6/14 (version 2.2) for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of any of my medical notes may be looked at by responsible individuals of the sponsor or research site and from regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐
4. I agree that my tissue samples (tumour sample and DNA extracted from my blood) will be sent to the Blizard Institute at Barts and the London School of Medicine where it will be anonymised and stored for the duration of the study, then stored at MEH Biobank. ☐
5. I agree to being recalled in the future (up to 5 years) if you need further clinical information and Biological samples from me. I understand that this part of the study is entirely optional. ☐
6. I agree to take part in the above study. ☐

Name of Patient	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

Version 2.2, 1/6/14
NRES Committee North West 14/NW/1080

Figure 2.1 Consent form to recruit patients for the study.

Retrospective approval

Archival tissue stored at the Moorfields Biobank was given approval for use by the Biobank REC under the same prospective reference (10/H0106/57-2012ETR28).

2.1.2 Funding and Sponsorship

Fight For Sight UK

Fight For Sight (FFS) is an eye charity that awarded me with a clinical fellowship to carry out the three year PhD with a grant of £189,957 (reference number 1327/28). Their ethos is to prevent and treat eye conditions. Part of the fellowship involved patient and public engagement in order to promote their research portfolio. This included carrying out lectures for the speakers network designed to engage with not only eye professionals, but also allied organisations including those involved in the eye trade (selling glasses/contact lenses) and those in non-eye related primary care. In addition, the FFS website provides a wealth of educational information on eye conditions accessible to the public.

National Institute for Health Research Clinical Research

The National Institute for Health Research (NIHR) clinical research network (CRN) comprises of 15 local clinical research networks working together to record and support clinical studies. The study was registered onto the national portfolio of studies in an anonymised fashion. This is accessible to members of the public who can learn about the project and see how many people have been recruited. Access to the portfolio is via the following link: www.crn.nihr.ac.uk/can-help/funders-academics/nihrcrn-portfolio. Furthermore, NIHR Moorfields Biomedical Research Centre provided an award amounting to £18,120 to help with consumable costs during the final PhD year.

Sponsor

Queen Mary University of London (QMUL) was the acting Sponsor of the study as defined in the Research Governance Framework for Health and Social Care Act 2005 and filed under ReDa reference 008621 QM. Furthermore, the study was insured by Arthur J Gallagher International through QMUL as the sponsor.

2.1.3 Informed Consent

Informed consent was obtained from all patients with the help of a specially designed peer-reviewed patient information sheet (Figure 2.2). The inclusion criteria for patients into the study included those aged between 18-90; both male and female; tissue diagnosis of SGC, mBCC or nodBCC using predetermined definitions as laid out in the RCPath datasets; first presentation; previous non-cancerous surgery to eyelid. Exclusion criteria included previous chemotherapy or radiotherapy; recurrence; children under 18; prisoners; young offenders.


KING'S HEALTH PARTNE

Guy's and St Thomas' NHS
NHS Foundation Trust

Queen Mary University of London

Moorfields Eye Hospital NHS
NHS Foundation Trust

St John's Institute of Dermatology
Guy's Hospital
Great Maze Pond
London SE1 9RT

Telephone: 020 718 86273 / 86277 / 89142

Fax 020 718 86113

Email: stru@gstt.nhs.uk

SUBJECT INFORMATION SHEET and INFORMED CONSENT FORM
Title of Study: Identification of genetic factors involved in morphoeic basal cell (mBCC) and sebaceous gland carcinoma (SGC) of human eyelid tumours with a view to identifying potential treatment targets.
Version and Date of Subject Information Sheet/Informed Consent: 3/7/14 Version 2.3
Ethics Committee Code: **14/NW/1080**

INTRODUCTION

You have been invited to participate in a study to look for the genes that may cause your eyelid tumour. Please take your time to decide whether you would like to participate. Discuss it with your friends and family. Please ask the study doctor or nurse to explain anything you do not understand.

WHAT IS THE PURPOSE OF THE STUDY?

This study is part of a PhD research project on eyelid cancer undertaken by Dr John Bladen (specialised registrar in Ophthalmology) and is designed to answer the following questions (which may or may not be relevant to you):

1) Why do some people develop eyelid cancer and some people don't?
2) Why do some people get more aggressive forms of cancer and some patients don't?
3) Why do some patients respond to their treatments and some don't?
4) Is there any way we can predict whether your eyelid tumour will get better or worse over time?
5) Can we develop tests to aid in the diagnosis of the aggressive cancer subtypes?
6) Can we use this information to develop better treatments?

WHAT WILL THE STUDY ENTAIL?

1) Donation of the removed cancer tissue sample that is taken as part of your normal treatment and routinely destroyed. Importantly, no extra tissue will be removed and your operation will not be modified in any way. In addition, it does not affect your diagnosis or subsequent treatment.
2) A member of the study team will look at your medical notes
3) A blood sample (approx 10mls or 2 teaspoons) to look at your DNA (genetic make up)

Figure 2.2 Patient information sheet. First page of the patient information sheet and informed consent form.

2.1.4 Sample retrieval

Material transfer and storage of tissue and blood

Both archival and fresh tissue samples, along with blood samples, were sent to Blizard Institute, Barts and the London School of Medicine, QMUL, 4 Newark Street, E1 2AT. Samples were be collected and personally couriered by myself and stored in the aforementioned Institute under the care of Professor Mike Philpott. At the end of the study, all samples were returned for storage at Moorfields Eye Hospital Biobank or destroyed.

Data collection, handling and record keeping

All tumour samples were anonymised at the point of retrieval. No patient identifiable details were collected. Anonymised data were kept on specific password protected, and encrypted computers. Excel software was used to collate results. In addition, the collated results were loaded onto the NIHR CRN portfolio as mentioned. Records will be kept for 5 years after completion of the project then erased. In addition, the project adhered to the Data Protection Act 1998 and Caldicott principles.

Fresh sample retrieval

Surplus fresh cancer tissue was removed, ensuring that the borders remained intact, from the patient prior to placement in formalin and embedded in paraffin for histological examination. During Mohs micrographic excision, the debulking tissue was taken for research. The research sample was placed into a nuclease free, UV treated 2 µl microcentrifuge tube. Marking of the skin surface during retrieval using a surgical marking pen allowed for correct tissue orientation for subsequent cryosectioning. Once the tube was labelled with the randomised number, the sample was placed into dry ice for snap freezing and transported to store in the -80°C freezer. Thickness of the sample rarely exceeded 5mm in the thinnest dimension and this allowed for successful rapid freezing using dry ice alone. An initial test comparison was made using liquid nitrogen for snap freezing and tissue integrity was similar in both as long as the sample remained small. At worst, a thick sample would show ice crystal formation with slight loss of tissue architecture on H+E staining, however, this did not affect the integrity of nucleic acid for extraction, including RNA. In addition, the use and transport of liquid nitrogen between tissue collection sites and the lab was prohibitive.

2.2 Preparation of tissue for nucleic acid extraction

2.2.1 Preparation and cutting of fresh tissue

Tissue was transferred to the cryostat-microtome (Leica CM3050S) machine on solid carbon dioxide (dry ice or Cardice®). Cutting of tissue was done at -23ºC (both the specimen and chamber temperature). Tissue was orientated so that transverse cross-sections were obtained with the epithelial surface on top and the subcutaneous layer at the bottom aided by the marking made at the time of surgery. Once correctly orientated, the tissue was embedded into a bead of optimal cutting temperature medium (OCT: Tissue-Tek® Sakura® polyethylene glycol and polyvinyl alcohol <11%) for DNA extraction or 0.1% DiEthylPyroCarbonate (DEPC, Sigma-Aldrich®) water for RNA placed on a chuck at -23°C. Tissue was cut at a thickness of either 10 µm onto a laser capture microdissection slide for nucleic acid extraction or 7 µm onto a plain glass slide (Shandon Colorfrost™ Plus, Thermo scientific) for histology.

2.2.2 Preparation and cutting of formalin fixed paraffin embedded (FFPE)

At the time of tissue collection, the rest of the sample was sent for histological analysis to ensure the borders are clear by fixing the tissue in formalin. The specimen was taken to the Institute of Ophthalmology (IoO) where it was cut and embedded into paraffin. Borders were then reported on to determine clearance for the patient. Once the clinical examination was completed and specimen reported on, the surplus FFPE tissue blocks were used for research purposes. Excess paraffin was trimmed off and research sections were cut 10 µm thick using a clean blade rinsed in DNA*Zap*™ (Life technologies™) and RNase*Zap*® (Applied Biosystems®, Life technologies™) on a Leica microtome RM2255. The blocked was cooled to aid cutting and prevent shredding. Tissue sections were floated on 0.1% DEPC water (Sigma-Aldrich®).

2.2.3 Laser capture microdissection of fresh tissue

Slide pretreatment

Tissue was placed onto a laser capture microdissection (LCM) slide covered with polyethylene naphthalate (PEN)-membrane (Carl Zeiss MembraneSlide 1.0 PEN) and

pretreated with UV light at 254 nm for 45 minutes. The UV rays convert and the relatively hydrophobic PEN membrane into a more hydrophilic structure that allowed for better adherence of the sections to the slide. Furthermore, the UV lights breaks down any potentially contaminating nucleic acids that may be on the surface of the slide. Tissue of 10 µm thick was placed onto the slide using the cryostat-microtome (Leica CM3050S) and the box immediately put onto dry ice.

Staining of nuclei using methyl green to determine area of interest

The LCM slides were removed from the box on dry ice and allowed to defrost briefly. When the tissue turn from opaque (frozen) to clear (just defrosted) a drop of methyl green was placed onto the tissue and washed off using DNase, RNase, and nuclease-free, deionised water (Qiagen® nuclease-free water) after 5 seconds only. The slide was air dried momentarily before being placed into the box on dry ice. Despite the rapid staining, this was more than adequate to visualize tumour tissue from normal structures (Figure 2.3). The glass bottle used to make the meythl green was cleaned using DNA*Zap*™ (Life technologies™) and RNase*Zap*® (Applied Biosystems®) Moreover, the Qiagen® nuclease-free water was used to make up the solution.

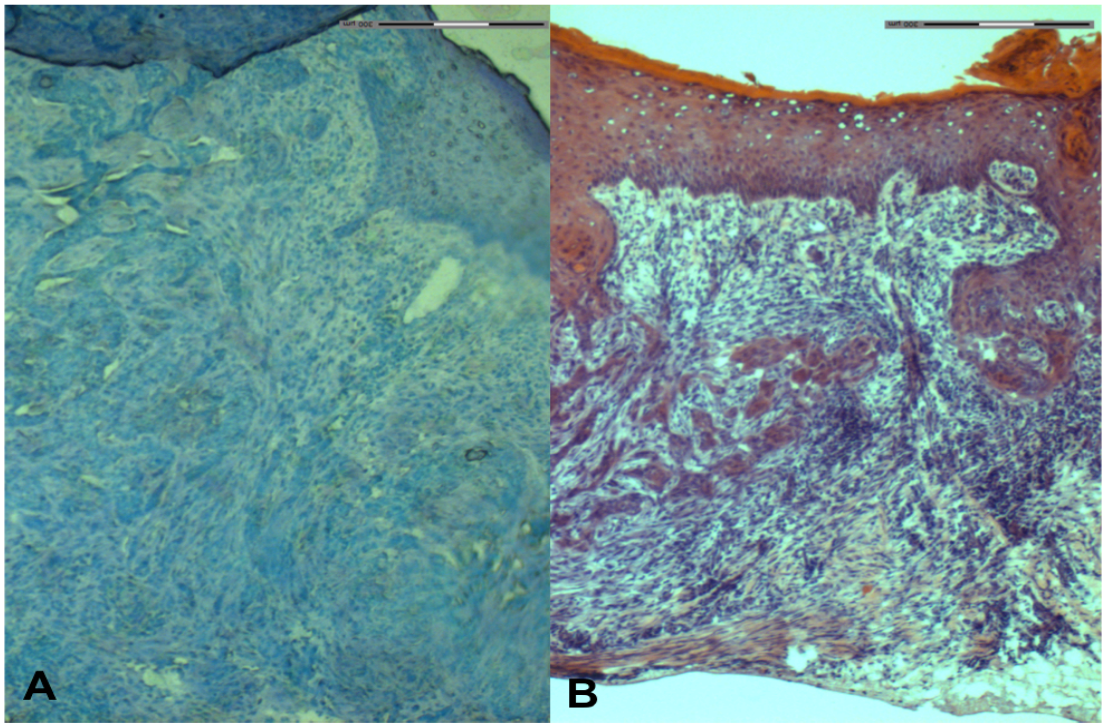


Figure 2.3 Immunostaining of mBCC. Immunohistochemistry with (A) Methyl green and (a) Fast Haematoxylin

P.A.L.M Microlaser

The PALM microBeam (version-1105) Carl Zeiss LCM device was used to cut and collect the specimens. The machine was cleaned before use to avoid contamination and one slide at a time was used with the rest kept cool in a box containing dry ice. The settings for LCM were speed at 15%, focus 79 with delta 2 at x10 magnification and cut energy of 35 J. At x5 magnification, cut speed was 30%, focus at 78 with delta 2 and cut energy was at 53 J. Two techniques were employed for removal of tissue from the slide depending on the size of the material of interest. Sizes smaller than 1mm utilized the laser pressure catapulting (LPC) mode where the circumference is ablated and the cut area subsequently catapulted into an opaque AdhesiveCap (Carl Zeiss). Tissue greater than 1mm was removed using two sterile, 25 gauge needles and placed into a 1.5 µl microcentrifuge tube containing 180 µl of buffer ALT and 30 µL of proteinase K (Qiagen® QIAamp® DNA micro kit). Tissue collected in the AdhesiveCap (Carl Zeiss) was inverted into 40 µl buffer ALT containg 20 µl proteinase K (Qiagen® QIAamp® DNA micro kit).

2.2.4 Laser capture microdissection of formalin fixed paraffin embedded (FFPE)

LCM slides were pretreated as per fresh tissue, but no staining with methyl green was carried out. Comparison with haematoxylin and eosin (H&E) staining of a preceding section at the time of laser allowed for accurate excision of target tissue (Figure 2.4). The PALM Microlaser settings for FFPE LCM were speed at 25%, focus 79 with delta 2 at x5 magnification and cut energy of 50 J. The two tissue removal techniques as described above was used and the sample placed into 320 µl of deparaffinization solution (Qiagen®).

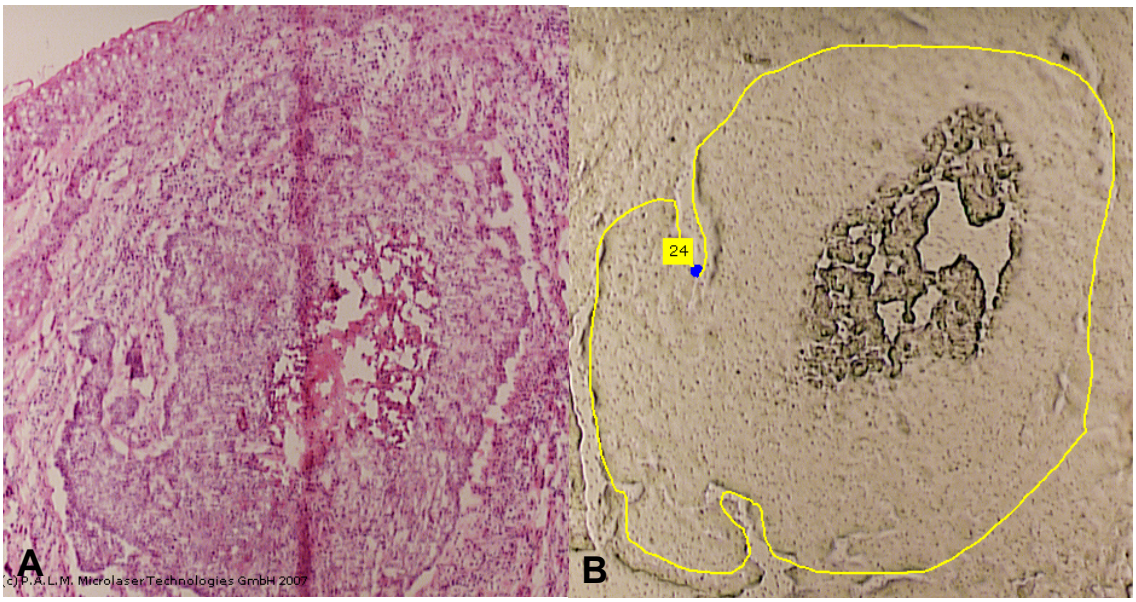


Figure 2.4 Laser Capture Microdissection (LCM) (A) H&E template used for LCM that is not stained (B) LCM of FFPE sample.

2.3 Extraction of nucleic acids

2.3.1 Total DNA extraction from fresh tissue

Tissue samples were placed into buffer ALT with proteinase K (Qiagen® QIAamp® Micro kit) with the amounts depending on tissue size as mentioned above, and mixed with pulse vortex for 15 seconds. The mixture was then incubated overnight for a minimum of 15 hours at 56ºC for efficient lysis. Subsequently, 200 µl AL buffer was added and pulse vortexed for 15 seconds, then 200 µl absolute ethanol put in, pulse vortexed for 15 seconds and incubated for 5 minutes at room temperature followed by 15 seconds of centrifuge at 6000g. The entire lysate was spun through a QIAamp® MinElute column (which had been stored at 4ºC prior to use) for 1 minute at 6000g. A new collection tube (CT) was used and 500 µl buffer AW1 added, and then spun for 1 min at 6000g. A clean CT was placed and 500 µl buffer AW2 placed, and then spun for 1 min at 6000g. Using a clean CT, the QIAamp MinElute was spun at full speed for 3 minutes to dry the silica membrane. Two elutions were carried out, both with 50 µl of buffer AE and incubated for 5 minutes. Finally, the column was spun to elute into a 1.5 µl microcentrifuge tube and stored at -20ºC.

2.3.2 Total DNA extraction from FFPE tissue

Tissue was placed initially placed into 320 µl of deparaffinization solution (Qiagen®), pulse vortexed for 10 seconds and briefly centrifuge for 15 seconds at 6000g. The mixed solution was incubated on a heat block at 56ºC for 3 minutes. On cooling to room temperature, the mixture was observed to see if it remains as a solution or solidifies. If the latter occurred, more deparaffinization solution (Qiagen®) was added in aliquots of 160 µl, the combination reheated to 56ºC for 3 minutes then cooled until the mixture remains a solution on cooling. Buffer ATL was added (180 µl), pulse vortexed for 15 seconds and centrifuged for 1 min at 11,000g to form a clear lower phase and murky supernatant (Qiagen® QIAamp® DNA FFPE Tissue kit). Thirty microlitres of proteinase K was added to the clear lower phase and mixed by pipetting only, then incubated overnight (minimum 15 hours) at 56ºC to ensure complete lysis. Partial reversal of the formaldehyde modification of nucleic acids was attempted by heating to 90ºC for 1 hour. Droplets formed on the top of the lid were removed by spinning at 6000g for 30 seconds. The clear lower phase was transferred to a new 1.5 microcentrifuge tube, allowed to cool to room temperature and 2 µl of RNase A was added for 2 minutes. Subsequently, 200 µl AL buffer was added and pulse vortexed for 15 seconds, then 200 µl absolute ethanol put in, pulse vortexed for 15 seconds and incubated for 5 minutes at room temperature followed by 15 seconds of centrifuge at 6000g. The entire lysate was spun through a QIAamp® MinElute column (which had been stored at 4ºC prior to use) for 1 minute at 6000g. A new collection tube (CT) was used and 500 µl buffer AW1 added, and then spun for 1 min at 6000g. A clean CT was placed and 500 µl buffer AW2 placed, and then spun for 1 min at 6000g. Using a clean CT, the QIAamp® MinElute was spun at full speed for 3 minutes to dry the silica membrane. Two elutions were carried out, both with 50 µl of buffer AE and incubated for 5 minutes. Finally, the column is spun to elute into a 1.5 µl microcentrifuge tube and stored at -20ºC.

2.3.3 Total DNA extraction from blood

Fresh blood was drawn from the antecubital fossa, and subsequently the DNA extracted within hours of removal by utilising the Qiagen® DNeasy® Blood and Tissue kit, and excess blood stored at -80ºC. Thirty microlitres of proteinase K was placed into a 1.5 microcentrifuge and 200 µl of whole blood added along with 200 µl buffer AL, followed

by pulse vortexed for 15 seconds. Incubation occurred for 10 minutes at 56ºC and the lysate spun to removed droplets for the top of the lid before adding 200 µl of absolute ethanol, pulse votexed for 15 seconds and placed into a DNeasy® Mini spin column. The mixture was spun for 1 minute at 6000g, follow-through discarded and clean CT place. Five hundred microlitres of buffer AW1 was placed and centrifuged for 1 minute at 6000g. Subsequently, 500 µl of buffer AW2 was put in and spun at full speed for 3 mintues. The CT was replaced with a clean one and spun for a further minute at full speed to prevent buffer AW2 carry over. Two elution's were carried out using 200 µl of AE buffer into a 1.5 microcentrifuge tube and incubation times of 5 minutes. The resultant elute was stored at -20ºC.

2.3.4 Determination of DNA fragmentation

Gel electrophoresis was used to run to determine the integrity of the DNA extracted. Agarose gel at 1% was made (0.3 g agarose dissolved in 30 ml of x1 TAE buffer) and set with 15 µl of GelRed™ nucleic acid gel stain (Biotium) and a comb placed at one end. The solidified gel was subsequently placed in 1X TAE running buffer. Loading buffer (1 µl) was mixed with 5 µl of DNA (at less than 0.5 ng/µl per lane) and placed into the wells. A 1kb Plus DNA ladder (invitrogen™ by Life technologies™) was inserted as a reference. The electophoresis was set at 100 mV for 45 minutes with the negatively charged DNA moving towards the positive electrode. Intact, high quality DNA was defined as large fragments greater than 50 kb (Fig 2.5). Purified DNA or cDNA should run as a single band.

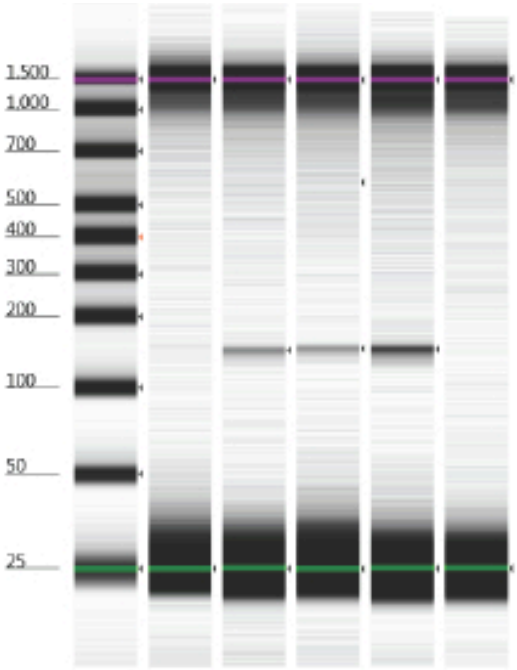


Figure 2.5 DNA electrophoresis. Purified DNA is run to see if it runs as a single band on gel electrophoresis. Lane one is the ladder. No bands are detected on lane 2 and 5. Lane 4 contains the highest concentration of DNA.

2.3.5 Total RNA including miRNA extraction from fresh tissue

RNA was extracted from fresh, LCM tissue using a modified Qiagen® RNeasy® Micro kit protocol. Tissue was placed into a 1.5 microcentrifuge tube containing 350 µl buffer RLT and 3.5 µl of β-mercaptoethanol (Sigma-Aldrich®), then pulse vortexed for 30 seconds. Three hundred and fifty microlitres of 70% ethanol, mixed with moderate pipetting for 15 seconds and the entire sample transferred into a RNeasy® MinElute spin column (which had been stored at 4ºC prior to use). Following, the column was spun at 8000g for 15 seconds, follow-through discarded and CT reused. Buffer RW1 was added at a volume of 350 µl, column spun at 8000g, the follow through discarded and CT reused. On column digestion of DNA occurred using 10 µl DNase1 solution mixed (by tube inversion only) with 70 µl Buffer RDD for 15 minutes at room temperature. A further wash with 350 µl Buffer RW1 occurred, centrifuging for 15 seconds at 8000g and discarding both the follow-through and CT. The column was placed into a new CT, 500 µl Buffer RPE added, spun for 15 seconds at 8000g and follow-through discarded, but the CT reused. Subsequently, the column was washed with 80% ethanol and spun for 2 minutes at 8000g. With the follow-through and CT discarded, the column was dried by

spinning for 5 minutes at full speed. Two elutions were carried out by placing 15 µl of nuclease free water directly onto the membrane and incubating for 5minutes before spinning at full speed for 1 minute into a 1.5 microcentrifuge tube. The resultant RNA was stored at -80ºC.

2.3.6 Total RNA including miRNA extraction from formalin fixed paraffin embedded tissue

Tissue was placed initially placed into a 2 µl microcentrifuge tube containing 320 µl of deparaffinization solution (Qiagen®), pulse vortexed for 10 seconds and briefly centrifuge for 15 seconds at 6000g. The mixed solution was incubated on a heat block at 56ºC for 3 minutes. On cooling to room temperature, the mixture was observed to see if it remains as a solution or solidifies. If the latter occurred, more deparaffinization solution (Qiagen®) was added in aliquots of 160 µl, the combination reheated to 56ºC for 3 mintues until the mixture remains a solution on cooling. RNeasy® FFPE kit from Qiagen® was used to extract the RNA. Buffer PKD was added at 240 µl volume and pulse vortexed for 15 seconds, and then centrifuged for 1 minute at 11,000g. Twenty microlitres of proteinase K was added to the lower phase and mixed by gentle pipetting. Initial incubation occurred at 56ºC for 15 minutes for proteinase K tissue digestion, followed by immediate incubation at 80ºC to partially reverse the formaldehyde induced cross links. The lower phase was transferred into a new 2 µl microcentrifuge tube and placed onto ice for 3 minutes before centrifuging for 15 minutes at full speed to create a pellet of insoluble tissue debris including crosslinked DNA. The supernatant was removed without disturbing the pellet and placed into a new 2 µl microcentrifuge tube. DNA digestion occurred with 25 µl of DNase booster buffer and 10 µl DNAse1 for 15 minutes at room temperature. Binding conditions were improved by adding 500 µl Buffer RBC, pulse vortexed and 1200 µl of absolute ethanol added using pipetting, then transferred into a RNeasy® MinElute spin column. The column was spun for 15 seconds at 8000g. The flow-through was discarded and CT reused. Buffer RPE was subsequently washed through the column twice, firstly 500 µl which is spun for 15 seconds at 8000g and secondly, 500 µl centrifuged for 2 minutes. Next, the RNeasy® MinElute spin column was dried by centrifuging for 5 minutes at full speed. Following, two elutions were carried out using 15 µl of nuclease free water and the resultant RNA stored at -80ºC.

2.3.7 Determination of nucleic acid concentration and purity

To ascertain larger quantities of nucleic acid, the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, Inc) was used to determine the concentration in ng/µl by testing 1.5 µl of undiluted sample. Although the concentration detection range of the spectrophotometer is from 2-3700 ng/µl for double stranded DNA and 2-3000 ng/µl RNA, any sample measuring less than 10 ng/µl were analysed using the Qubit® 2.0 Fluorometer (invitrogen™ by Life technologies™). The spectrophotometer was set to nucleic acid application mode and to either DNA-50 (double stranded DNA) or RNA-40 (total RNA). Up to 1.5 µl of sample was placed onto the unit and on activating the system, an absorbance curve was produced along with the concentration and purity ratios. A target purity ratio relating to protein contamination of 260 nm to 280 nm (A_{260}/A_{280}) used for DNA was ~1.8 and ~2.0 for RNA. A target purity ratio relating to organic compound contamination of 260 nm to 230 nm (A_{260}/A_{230}) was >2.0. For smaller concentrations the spectrophotometer became inaccurate and a benchtop fluorometer, Qubit® 2.0 Fluorometer (invitrogen™ by Life technologies™), was used in conjunction with the Qubit™ high sensitivity assay kit (DNA range 10 pg/µl to 100 ng/µl and RNA range 250 pg/µl to 100 ng/µl). A working solution is made using 1 µl of fluorescent reagent to 199 µl of buffer per sample including two standards. Thin walled 0.5 µl PCR tubes were used for accurate readings. The working solution and the test sample were vortexed and readings were taken after 2 minutes of incubation.

2.3.8 Determination of RNA Integrity Number (RIN)

Total RNA underwent RIN analysis using an Agilent 2100 bioanalyzer instrument (Agilent technologies, Inc). LCM samples yielded very low concentrations of RNA, hence an Agilent RNA 6000 pico kit was employed whose range was 50-5000 pg/µl. The electropherogram was observed to look at 18S and 28S fragments along with the presence or absence of degradation products (Figure 2.5). RIN values greater than 7 were used for RNA sequencing.

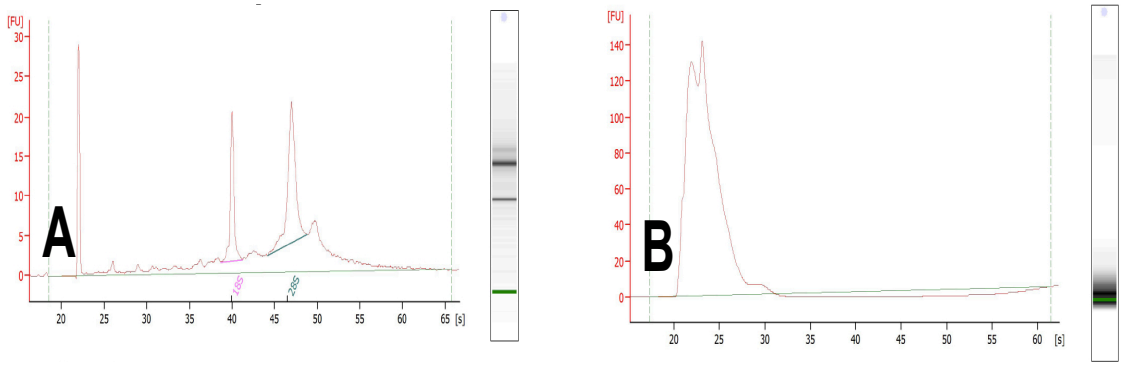


Figure 2.6 RNA RIN Electropherograms (A) RNA extracted from fresh tissue with a RIN of 8.6 and (B) RNA extracted from FFPE with a lower RIN of 2.6.

2.4 Polymerase chain reaction (PCR)

2.4.1 PCR primer design

Primers were devised against cDNA sequence taken from university of California Santa Cruz (UCSC) using GRCh38 and designed to be 18-30 nucleotides in length, have similar melting temperatures (T_m), a GC content of 50-60% where possible, and calculated using the following formula; $T_m = 2°C \times (A+T) + 4°C \times (G+C)$. In order to avoid amplification from any contaminating genomic DNA, primers were constructed to span exons. Furthermore, the primer pair was separated by an intron between 1000-1000,000 bp long with the final product length between 90 to 200 bp and any unintended products being greater than 1000bp as confirmed by primer basic local alignment search tool (BLAST) from the National Centre for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/tools/primer-blast). Possible primers were crosschecked against primer design websites Primer3web version 4.0 (http://primer3.ut.ee/) and the primer basic local alignment search tool (BLAST). Specificity of primers was checked again using the BLAST from Ensembl (http://www.ensembl.org/Multi/blastview). Moreover, OligoAnalyser3.1 was used to analyse homo- and hetero-dimer potential. The primer sequences designed for use are summarised in Table 2.1.

Customised oligonucleotides for micoRNA were designed (Table 2.2) in conjunction with Invitrogen™ (Life technologies™) to use Taqman advanced technology qPCR. This provided a better technique for short such short lengths of DNA.

Gene	Forward (5'>3')	Length	Tm	GC	Self	Self 3'
	Reverse (5'>3')			%	comp	comp
DDR1	GAGCGATGAGAGGTGTCTGAA	21	59.53	52.38	3	1
	CATAGCTCCTGATCCCTCGG	20	59.11	60.00	4	2
PDGFA	TCCGTAGGGAGTGAGGATTCTT	22	60.02	50.00	6	3
	AATGACCGTCCTGGTCTTGC	20	60.32	55.00	5	2

Table 2.1 Primer design for mRNA. Primers designed for use in real time quantitative PCR from messenger RNA derived cDNA; Tm, melting temperature; self comp, self complimentary binding.

Mature miRNA	Accession number	TaqMan oligo
hsa-miR-16-5-p	MIMAT0000069	UAGCAGCACGUAAAUAUUGGCG
hsa-miR-34a-5p	MIMAT0000255	UGGCAGUGUCUUAGCUGGUUGU
hsa-miR-205-5p	MIMAT0000266	UCCUUCAUUCCACCGGAGUCUG
hsa-miR-150-5p	MIMAT0000451	UCUCCCAACCCUUGUACCAGUG
hsa-miR-199a-3p	MIMAT0000232	ACAGUAGUCUGCACAUUGGUUA
miR-26a-5p (control)	MIMAT0000082	UUCAAGUAAUCCAGGAUAGGCU

Table 2.2 Primer design for microRNA. Customised primers for Taqman® real time quantitative PCR from microRNA derived cDNA. MiRNA were chosen from Nanostring® nCounter® human v2 MicroRNA expression assay.

2.4.2 Reverse transcriptase polymerase chain reaction (RT-PCR)

Messenger RNA to copy DNA (cDNA)

Input RNA for reverse transcription (RT) had been treated with DNase 1 to ensure minimal genomic DNA was present in the sample. Concentrations of RNA were within the recommended range of 2 – 200 ng/µl and dissolved in nuclease/RNase free water. Quality nondenatured RNA was used as determined by the Agilent 2100 bioanalyzer instrument (Agilent technologies, Inc). SuperScript™ III First-Strand Synthesis SuperMix (invitrogen™ by Life technologies™) was used for cDNA synthesis. It also contains RNaseOUT™ to protect against degradation of RNA input template. Reaction mixture consisted of 2 µl of RT enzyme mix, 10 µl 2X RT reaction mix, variable volume of input RNA depending on concentration and made up to 20 µl with DEPC-treated water in a PCR tube. The tube was mixed gently and heated to 25ºC for 10 minutes followed by 30 minutes at 50ºC. Termination of the reaction occurred at 85ºC for 5 minutes then the tube was placed on ice. Two units of RNase H (1 µl) from *E.Coli* was added to remove RNA template from the RNA:cDNA hybrid molecule after first strand synthesis and incubated for 20 minutes at 37ºC. Template cDNA quality was examined using

spectrophotometry and agarose gel electrophoresis as described in 2.3.4 and 2.3.8 respectively.

Copy DNA library preparation from MicroRNA including amplification

A four step process (poly(A) tailing, ligation, reverse transcription and miR-amplification) was undertaken to create a library of cDNA from isolated miRNA (2.3.8) using TaqMan® Advanced miRNA cDNA synthesis kit. An input range of 1-10 ng of total RNA containing enriched miRNA was used. Poly(A) tailing cocktail of 0.5 µl 10X poly(A) buffer, 0.5 µl ATP, 0.3 µl poly A enzyme, 1.7 µl RNase-free water was prepared in a 1.5 ml microcentrifuge tube and combined with 2 µl of input RNA by vortexing and centrifuged to remove bubbles. The mixture was placed into a thermal cycler and incubated at 37°C for 45 mins of polyadenylation followed by 65°C for 10 mins to stop the reaction. A ligation cocktail of 3 µl 5X DNA ligase buffer, 1.5 µl RNA ligase, 4.5 µl 50% PEG 8000, 0.6 µl 25X ligation adapter, 0.4 µl RNAse-free water was made and after mixing, all 10 µl was added to the poly(A) tailing reaction. The vortexed mixture was placed into a thermal cycler and incubated at 16°C for 60 mins of ligation. A RT cocktail of 6 µl 5X RT buffer, 1.2 µl dNTP (25 mM each), 20X universal RT primer, 3 µl 10XRT enzyme, 3.3 µl RNase-free water was made and after mixing, all 15 µl was added to the ligation product. The vortexed mixture was placed into a thermal cycler and incubated at 42°C for 15 mins of reverse transcription followed by 85°C for 5 mins to stop the reaction. An amplification cocktail of 25 µl 2X miR-Amp master mix, 2.5 µl 20X miR-Amp primer mix, 17.5 µl RNase-free water was made and after mixing, all 45 µl was added to a 5 µl aliquot of RT product. The rest of the RT product was stored for up to 2 months at -20°C. The amplification mixture was run according to conditions in Table 2.1. The product was then immediately used for qPCR or stored at -20°C.

Step	Temperature	Duration	Cycle
Enzyme activation	95°C	5 min	1
Denaturation	95°C	3 sec	
Annealing/Extension	60°C	30 sec	X14
Stop	99°C	20 min	1
Hold	4°C	hold	1

Table 2.3 Step 4 microRNA to cDNA formation. Amplification thermocycling conditions as part of the forth step in miRNA RT cDNA formation.

2.4.3 Real time quantitative PCR (RL-qPCR)

Two detection systems were used to carry out RL-qPCR; Probe based (TaqMan® Assay, Life technologies™) and dye based (SYBR®, invitrogen™ by Life technologies™, Thermofisher) assays. The Applied Biosystems 7500 RT PCR system was used for both RL-qPCR techniques.

Probe based assay

High specificity minor-groove binding (MGB) Taqman® probes were used for genetic validation of the miRNA data. The main advantage is the enhanced specificity from the custom probe designed against the exact sequence. A 1:10 dilution of cDNA was made by combining 5 µl miR-Amp amplification product with 45 µl 0.1 TE buffer. A cocktail of 10 µl 2X Fast advanced mastermix, 1 µl 20X Taqman® advanced miRNA assay, 4 µl RNase-free water were vortexed and placed into a PCR well. Subsequently, 5 µl of diluted cDNA template was added to each well and sealed with an adhesive cover. In addition, an endogenous control gene, hsa-miR-26a-5p, was chosen as it is relatively constant and abundant within tissues. A minimum of three replicates for each reaction was performed and the endogenous control placed on every well appropriate intraplate analysis. No sample library and PCR mix control were also on each plate. Normalisation of the gene of interest was performed against this control gene in a similar fashion to mRNA analysis. Thermocycling settings are set out in Table 2.4.

Step	Temperature	Duration	Cycle
Enzyme activation	95°C	20 sec	1
Denaturation	95°C	3 sec	
Annealing/Extension	60°C	30 sec	X40
Hold	4°C	hold	1

Table 2.4 Taqman® RL-qPCR settings. Taqman® advanced miRNA assay settings using the Applied Biosystems 7500 RT PCR system.

Analysis involved the use of ExpressionSuite™ Software to view the amplification plots and set both the baseline and threshold values. Calculation of the relative expression used the comparative formula $\chi=2^{-\Delta\Delta Ct}$ where χ represents the fold difference and Ct represents the average threshold cycle of the triplicate samples. ΔCt refers to the difference between the gene of interest and the housekeeping gene hsa-miR-361-5p. $\Delta\Delta Ct$

represents the difference between the ΔCt cDNA of the tumour compared to the ΔCt cDNA of the tarsal plate.

Dye based assay

SYBR® green dye-based assay was used to analyse mRNA extracted from tumour and surrounding reference stroma. A minimum of three replicates of each reaction was performed and the standard thermocycling mode was used. For volumes of 20 µl per well, the plate contained 10 µl SYBR® Select mastermix 2x, variable concentration of forward and reverse primers (<200nM in standard thermocycling mode) and variable concentration of cDNA template in RNAse free water. The mixture was centrifuged removed air bubbles and bring the contents down. A standard thermocycling reaction was setup up and individualise by matching the annealing temperature to the melting points (T_m) of the primers (Table 2.6). Negative control sample omitted the template cDNA. A post-amplification dissociation (melt) curve was performed to detect primer-dimer artefact and reaction specificity by observing a single peak. Amplification plots were analysed to determine the threshold cycles (C_T) for each curve by setting the threshold above the baseline, background fluorescence signal but within the exponential growth phase of the amplification curve. The relative mRNA expression values were calculated using the formula $\chi=2^{-\Delta\Delta Ct}$ where χ represents the fold difference and Ct represents the average threshold cycle of the triplicate samples. ΔCt refers to the difference between the gene of interest and the housekeeping gene GAPDH. ΔΔCt represents the difference between the ΔCt cDNA of the tumour compared to the ΔCt cDNA stroma.

	Primer $T_m \geq 60°C$			**Primer $T_m < 60°C$**		
Step	**Temperature**	**Duration**	**Cycles**	**Temperature**	**Duration**	**Cycles**
UDG activation	50°C	2 min		50°C	2 min	
DNA Poly	95°C	2 min		95°C	2 min	
Denaturation	95°C	15 sec		95°C	15 sec	
Annealing	60°C	1 min	X40	50-60°C	15 sec	X40
Extension	60°C with annealing			72°C	1 min	
End and dissociation						

Table 2.5 Standard RL-qPCR. Standard cycling mode for mRNA derived cDNA, real time quantitative PCR according to primer T_m less or greater than 60°C. DNA poly, DNA polymerase activation.

2.5 Nucleic acid decipherment

2.5.1 Whole exome sequencing

Next generation whole exome sequencing was carried out on DNA extracted from fresh tumour samples. In addition, blood was extracted from the same patient and sent for whole exome sequencing for paired tumour-normal (blood) analysis. Gene encoding regions covering 1.5% of the human genome are sequenced. A minimum of 1.5 µg of DNA was required for library preparation and determined using a Qubit® 2.0 Fluorometer (invitrogen™ by Life technologies™) which had the advantage of detecting double stranded DNA. Intact, high quality DNA (defined as large fragments greater 50 kb) was confirmed using agarose gel electrophoresis. Library preparation using Agilent SureSelect™ (51Mb) version 5 capture products and whole exome sequencing on an Illumina® HiSeq™2000 100bp paired end analysis with a minimum 100X mean coverage was carried out by Oxford Gene Technology (OGT). Output files in FASTQ format were processed through our bioinformatics pipeline (section 2.7.1).

2.5.2 RNA sequencing

Extraction of high quality RNA, as determined by the bioanalyzer ,was required for RNA sequencing. Due to the small sample sizes and meticoulous LCM, low concentrations of RNA were obtained for some samples. To allow for an impartial comparison, a low input library preparation protocol (SMARTer® Universal low input RNA kit from Clontech®) capable of a range between 200 pg to 10ng was employed for all samples regardless of total yield. Amplification of the low input samples only would have introduced bias when comparing against the standard library preparation. Despite the low input range, samples were greater than 100ng total yield. Moreover, RNa was poly-A tail selected. Synthesised and adapter ligated copy DNA (cDNA) was then sequenced using 100 bp paired-ended reads on an Illumina® HiSeq™2000 averaging 13 million paired end reads per sample (12 samples in 1 lane) by OGT. Output FASTQ files were then processed in out RNA sequencing cancer pipeline (section 2.7.2).

2.5.3 Nanostring® nCounter® human v2 MicroRNA expression assay for SGC RNA

Total RNA was extracted from FFPE tissue and a minimum of 100 ng in 10 µl was required for assay analysis. Total RNA typically comprises of only a small amount of micro RNA (miRNA) and varies between tissues. (Liang et al., 2007) Sensitivity of the hybridisation reaction was previously shown to be good for FFPE tissue and contains unique oligonucleotide tags for 800 human miRNA based on the miRBase version 18. Output files were represented in the nCounter analysis system to obtain expression levels and these results were put through our miRNA pipeline.

2.5.4 Affymetrix™GeneChip® Human Genome U133 Plus 2.0 Array for SGC RNA

FFPE total RNA from SGC samples along with tarsal plate controls from a different individual were sent for Affymetrix™Human Gene U133 plus 2.0 array analysis by the University College London (UCL) genome centre. The total number of transcripts covered were 47,000 representing 38,500 well characterised genes and 6500 new ones. Sequence clusters were created from the UniGene database (Build 133, 20th April, 2001) including the 'plus' element of 6500 genes (Build 159, 25th Jan, 2003). A minimum of 50 ng of total RNA was required for chip assessment.

2.6 Nucleic acid analysis

2.6.1 DNA cancer pipeline and bioinformatics

FASTQ files were received from OGT and underwent a specific validated pipeline in conjunction with the QMUL bioinformatics unit which has been published (Okosun et al., 2013). All data underwent three main stages, 1) pre-calling scrutiny, 2) somatic (cancer) variant calling and 3) copy number analysis (Figure 2.7). Pre-calling involved quality control of FASTQ files, followed by read alignment against human genome 19 (hg19) using the backward search with Burrows-Wheeler transform, a second generation alignment (BWA) tool.(Li and Durbin, 2009) The resultant sequence alignment map (SAM) output file was converted into a compressed binary version (BAM file) that underwent sorting and removal of PCR duplicate paired-end reads using the Picard tool (version 1.86). The latter was to prevent false positive single nucleotide variant (SNV) calls. As a consequence, a BAM file with mapped data and a HTML file with the number of reads discarded was produced. Local realignment around indels, recalibration of quality scores and base quality recalibration using Genome Analysis ToolKit (GATK) unified genotyper was carried out (version 2.3.9). Variant calling of single nucleotide variants (SNVs) and short insertions/deletions (indels) occurred with the genotyper. The output variant call format (vcf) file could then be viewed as text in an excel format (Microsoft) or an integrative genomics viewer (IGV, Broad Institute). Simultaneous analysis of tumour–germline (blood) paired sequence analysis was carried out using Strelka pipeline. This has been shown to be particular good in cancer where tissue samples may have low tumour cell purity (down to 5%). This provided a lower false positive variant number than doing an initial independent analysis of the tumor and germline data, followed by a comparison. A study by Cai et al compared the four most popular tumour-germline programs (namely Strelka, Varscan, Somaticsniper and Mutect2) and ran data on 32 WES samples: it gave Strelka and Mutect2 the highest score for sensitivity and accuracy.(Cai et al., 2016) Known limitations of the program is that it requires a matched normal sample, paired end reads and should be a minimum of half the tumour depth. A variant was only considered as a possible significant change if a read depth greater than 20 was seen in the tumour. Subsequently, an attempt to determine any potential biological action of SNVs and small indels occurred by undertaking variant annotation. This included variant location, effect, evolutionary conservation and

regulatory effect. Lastly, copy number analysis was analysed using three different approaches. ASCAT assessed the genotyping information of identified variants between normal and tumour. VarScan2 (version 2.3.5) and DNACopy (R Bioconductor package version 2.12) used the number of aligned reads for exons as markers between normal and tumour.(Koboldt et al., 2012) Consequently, gene annotation took place in conjunction with RNA data for the formulation of potential pathways or molecular milieu specific to the tumour types. WES data was deposited to the European Genome-phenome Archive underthe accession number EGAS00001001915.

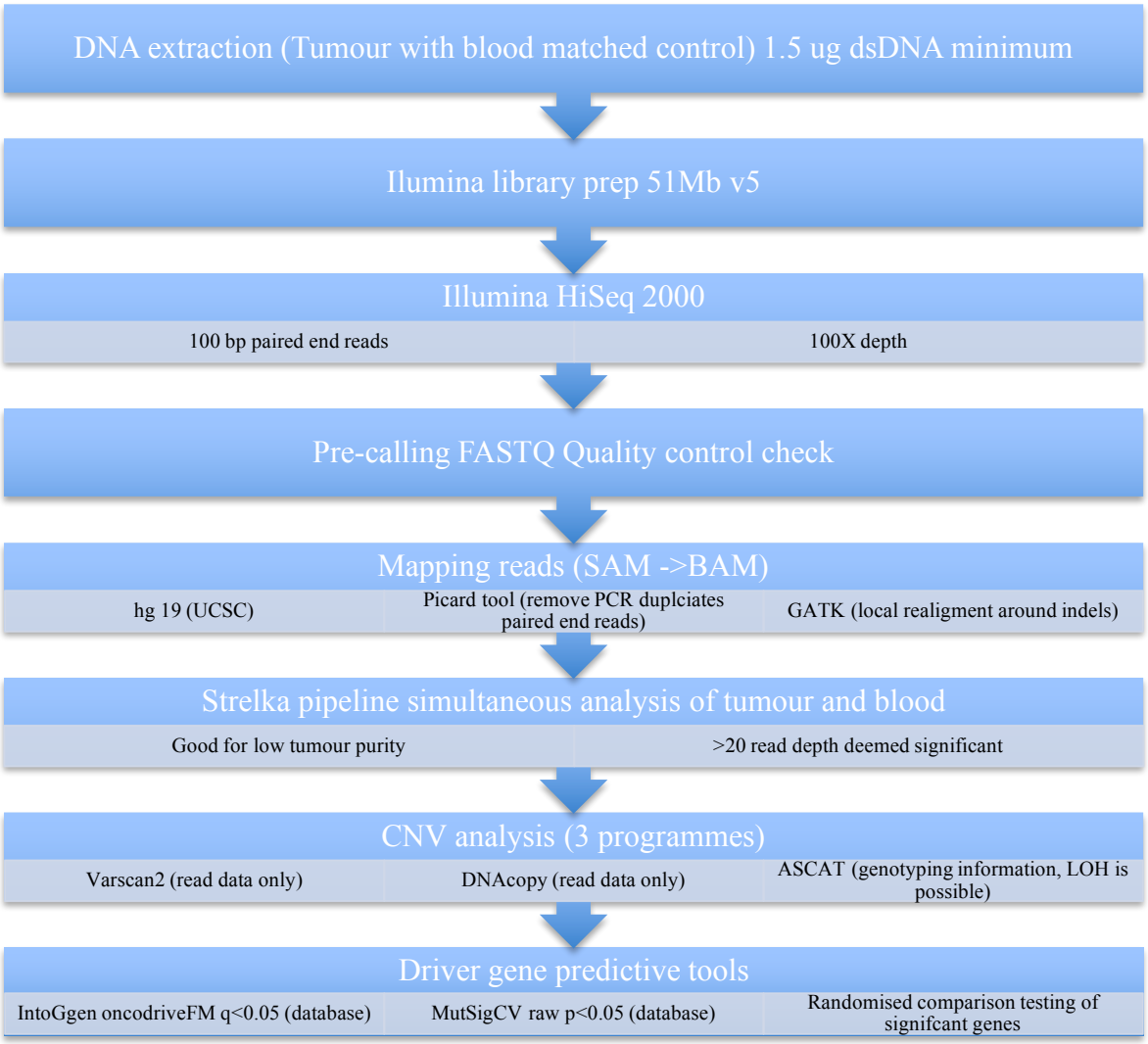


Figure 2.7 Whole exome sequencing pipeline. Process from DNA extraction to the determination of significant driver genes and copy number analysis. SAM, sequence alignment map; BAM, binary alignment map; GATK, genome analysis toolkit, LOH, loss of heterozygosity; CNV, copy number variation.

Predictive functional databases

Sorting intolerant from tolerant (SIFT)

Predicting the effects of non-synonymous variants found in coding regions of DNA on protein function.(Sim et al., 2012)

Polymorphism phenotyping version 2 (Polypen-2)

Another tool that predicts the functional effects of non synonymous human single nucleotide variants.(Adzhubei et al., 2010)

dbSNP

Single nucleotide polymorphisms are the most common genomic variation that can occur every 1200bp with the vast majority (99%) being a single nucleotide substitution.(Sherry et al., 2001) dbSNP is a database that catalogues genetic variation.

Catalogue of somatic mutations in cancer (COSMIC)

This database is continually updated with new somatic mutations in cancer and attempts to collate the data into a user friendly format and splits it into five sections.(Bamford et al., 2004, Forbes et al., 2011) Initially, it focused on somatic (cancer) point mutations within coding regions, however, with the advent of whole genome sequencing, non coding regions are included along with structural rearrangements and gene fusions. Access to the database is at http://www.sanger.ac.uk/cosmic. All mutations held in the database area assigned a Cosmic Somatic Mutation identifier (COSM id) to allow external database such as emsembl to access the correct record within COSMIC. Moreover, these COSM id remain as stable between the updating of the database. Further integration with external databases, acquisition of the huge sequencing data that is being carried out and curating individual driver genes as has been done with *TP53*.

DriverDB

The annotation databases mentioned above are incorporated with driver mutation bioinformatic algorithms driver and exome sequencing data to identify clinically significant changes in cancer.(Cheng et al., 2014)

Integrative Oncogenomics Cancer Browser (IntOGen)

This portal combines data from The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) and data from cancer journals. Currently, it contains data from 6792 WES samples from 48 projects and 28 cancer types (May 2016), but is constantly been updated. Driver mutations are placed into two broad categories, high-confidence driver (HCD) that is seen in the in the TCGA and candidate driver (CD) with an expectedly higher false-positive rate. The limitations is that it does not take into account the background mutation rate of the tumour and skin tumours including BCC have high levels which can increase the false positive for SNV calls. In addition, there are no WES samples relating to BCC which means that the driver mutation call are relating to different cancers, thus may not be as relevant and their significance is unknown for BCC.

Mutation significance (MutSig)

A software programme for the Broad Institute that was originally set up to look at background mutation rates (bmr) across the genome, it has evolved to analyse exome data to see if a gene is mutated more often than expected by chance.(Lawrence et al., 2013) It corrects for variation by using patient specific mutation frequency, spectrum and gene specific background mutation rates by incorporating expression levels and replication times. For example, melanoma has bmr of 100/megabase of DNA versus 5/megabase for a paediatric tumour and so corrects for the predicted noise. It takes into account the non-synonymous:synonymous SNV ratio for a particular gene, identifies hotspots within the samples given (10 mBCC) plus comparing to known hotspots using its database. The limitation is that it performs better with larger sample sizes and in this study the samples size is relatively small.

Identification of mutational driver genes and significantly mutated pathways

We used two methods to identify different positive selection signals that occurred in driver mutations: OncodriveFM(Gonzalez-Perez et al., 2012), implemented in the IntOGen software(Gonzalez-Perez et al., 2013) (Gonzalez-Perez et al. Nature Methods 2013), and MutSigCV(Lawrence et al., 2013) (Lawrence et al. Nature 2013). We applied the two algorithms in mBCC and nodBCC samples separately. For OncodriveFM, we used corrected *p*-values $q < 0.05$ for significance. For MutSigCV, due to the small sample size, *p*-values could not be adjusted, so raw $p < 0.05$ was applied. Overlapped genes that

were identified being significant by both methods were further selected. Significantly mutated pathways were also identified using OncodriveFM ($q < 0.05$). These identified pathways were further compared between mBCC and nodBCC groups.

Randomisation comparison test

To identify mBCC and nodBCC specific driver genes, we focused on genes significantly mutated in one subtype but not so much in the other based on OncodriveFM and MutSigCV results, e.g., genes within OncodriveFM $q < 0.05$ in one group but $q > 0.1$ in the other group. We then applied a randomisation procedure, where the pool of 20 total BCC samples was randomly split into two groups of equal size (n=10). We then applied OncodriveFM and MutSigCV on the two groups and recorded the significance values for our subtype specific candidate genes. We repeated the randomisation procedure 100 times and calculated the transformed p-values from OncodriveFM and MutSigCV, which generally followed a normal distribution for each gene. We next calculated the probability that the observed significance in mBCC or nodBCC group was stronger than that expected by chance for each candidate. Significant subtype genes from the randomisation test (two-tail $p < 0.05$) were further identified and regarded as the true mBCC or nodBCC specific genes, based on OncodriveFM and MutSigCV significance values, respectively. We also applied this procedure to identify specific mutated genes to eyelid nodBCC and nodBCC elsewhere.

2.6.2 RNA sequencing cancer pipeline

FASTQ output files underwent data quality check using FastQC, raw reads were aligned to the reference genome hg19 using Tophat2.(Kim et al., 2013a, Andrews) The number of uniquely aligned reads (mapping quality score $q > 10$) to the exonic region of each gene were counted using HTSeq based on the Ensembl annotation (version 75).(Anders et al., 2015, Cunningham et al., 2015) Genes that were expressed in a low level across all samples were excluded. For the 6 BCC tumour and stroma samples, only genes that achieved at least one count per million (cpm) mapped reads in at least 6 samples were kept, leading to 19,101 filtered genes in total. An abundance estimation was made looking at fragment per kilobase exon (FPKM). Normalisation for read counts took into account the length of the transcript, as longer transcript tend to be sequenced more than shorter transcripts. After applying scale normalization, read counts were converted to log2-cpm

using the voom function with associated weights ready for linear modeling.(Law et al., 2014) Differential expression analysis was then performed using the limma R package, using tumour-stroma paired analysis with an additive model design adjusting for baseline differences between patients.(Ritchie et al., 2015) A double threshold of raw p value <0.01 and an absolute fold change > 2 were used to define significantly differentially expressed (DE) genes. Gene Set Enrichment Analysis (GSEA) was performed for mBCC and nodBCC tumour-stroma comparisons using the GSEA tool to identify the dysregulated canonical pathways from the Molecular Signatures Database (MSigDB-C2, version 5.0) (Subramanian et al., 2005). Common and specific pathways between mBCC and nodBCC were further identified. GSEA performed for SGC used two pathway databases, KEGG and REACTOME in order to identify common pathways specific to SGC. RNAseq data has been deposited to the Gene Expression Omnibus using accession number GSE937997.

Heatmap representation of RNA seq data

Hierarchical clustering analysis was performed in R. This groups similar values into clusters then produces a dendogram that shows a hierarchy of clustering, but in an unsupervised way i.e. without taking into account the experimental variables such as tissue type. Samples were grouped into columns and gene expression into rows. These genes underwent a Z-score which soothes logFC, which allows event those genes with small changes to be seen on the plot. Z = (expression in tumour sample – mean expression in normal reference stromal tissue)/SD of expression in normal reference. A Z=0 means the gene expression is the same as the mean, or to put it another way the mean score for the gene becomes 0 with a maximum standard deviation=1. Thus, the gene expression in each sample is compared to 0 and is assigned a colour whether it is up (red) or down (blue) regulated. The viewer can interpret the colour as a standard deviation of the mean and has an intuitive idea of variation of that gene.s

Geneset enrichment analysis (GSEA)

This tool focuses on a geneset defined as a group of genes that share biological function, chromosome location or regulation). These genesets are taken from MSigDB-C2, KEGG and REACTOME which collects unbiased published expression profiles. Looking at the significance of a single gene may miss certain pathway activation especially where genes don’t reach significance due to mutational noise and having a list of significant genes

without biological theme can be confusing. Looking at a geneset and noticing just a small increase in expression (for example 10% increase in expression across many genes in the geneset) can be more significant that a large over expression of just one gene. Significance cut off include false discovery rate (FDR) q-value 0.10 (whereby 9 out of 10 times the result is valid), p-value <0.05 (enrichment score for a single gene set without taking into account the multiple hypothesis gene testing, taking into account geneset size and comparing against the empirical null distribution as FDR rate, thus the value is less than FDR). Thus, an enrichment score (ES) was calculated to see to which a geneset is correlated to ta particular pathway. ES was plotted against the gene list rank, with those closest to the y-axis more strongly correlated to the pathway. In addition, normalised ES was used to take into account for the difference in geneset size.

2.6.3 SGC MicroRNA cancer pipeline

MiRNA expression raw and normalised data were first generated from Nanostring nCounter Digital Analyzer. Filtered data were log2 transformed and further normalised using the quantile normalisation method in R, to make sure values across samples were comparable. DE analyses were performed using the limma R package.(Ritchie et al., 2015) The pairwise comparisons of nodular versus control and pagetoid versus control were first conducted, followed by the one-way ANOVA amongst the three groups. Common DE miRNAs between nodular and pagetoid against control, as well as unique DE miRNAs to the nodular and pagetoid groups specifically were identified.

2.6.4 Affymetrix™SGC RNA pipeline

Raw data from sebaceous gland carcinoma samples from the Affymetrix Human Gene 2.0 array were analysed using R. Expression data was normalised with guanine cytosine robust multiarray average (GCRMA) preprocessing algorithm which converts the raw expression data taking into account the GC content.(Wu Z, 2004) Data was then filtered using standard deviation to select the most variable probes and reduce the dimensionality. P values were adjusted using the Benjamini-Hochberg method to obtain a q-value.(Wu Z, 2004) DE probes between control samples and those taken from a pagetoid and between control samples and those taken from a nodular SGC were identified using the limma R package.

2.7 Protein expression

2.7.1 Immunohistochemistry

Haematoxylin and eosin (H&E)

H&E staining was performed using standard methods. FFPE tissue were dewaxed using xylene for two changes of 2 minutes each, alcohol for two changes of 2 minutes each and washed in tap water for 5 minutes. Subsequently, they were placed into haematoxylin for 5 minutes, followed by a wash with tap water for 3 minutes. Aqueous blueing solution (Scott's tap water substitute) was used followed by a further 3 minutes was in water. Two dips in 1% acid alcohol are performed and a ensuing wash in tap water for 3 minutes. Eosin staining of sections was performed for 5 minutes, a wash in tap water for 5 minutes and lastly a dehydrating step using increasing alcohol concentration. Fresh frozen tissue could be placed straight into haematoxylin for the allotted time, although a fast haematoxylin stain only was used for LCM orientation. Sections were sealed and covered using mounting medium (Sigma-Aldrich®) by hand or using the automated Thermo Clearview Coverslipper (Thermo Fisher Scientific, Inc).

3-3'-diaminobenzidine (DAB)

DAB immunostaining of 5 μm formalin-fixed, paraffin-embedded (FFPE) tissue sections was performed for the Hh pathway proteins including PTCH1, SMO, Gli1 and Gli2 (Abcam). Tissue sections were mounted on slides, deparaffinised in xylene then alcohol, followed by treatment with the DAKO EnVision™FLEX+ System (Dako). Antigen retrieval was performed using the DAKO PT LINK machine at 97°C for 20 minutes under alkaline (pH 9.0) or acidic (pH 6.0) conditions. Endogenous peroxidase activity was blocked using EnVision™FLEX peroxidase-blocking reagent. Tissue was incubated with primary antibody for 30 minutes, followed by incubation with a secondary antibody (Envision™FLEX/HRP LINKER; 20 mins) and then visualisation of the primary-secondary reaction with EnVision™FLEX substrate working solution containing buffer and DAB (10 mins). Positive control tissue included breast, brain, testes and intestine for PTCH1, SMO, Gli1 and Gli2 respectively. Negative controls involved all steps excluding the primary antibody. Finally, sections were counterstained with haematoxylin using the Gemini autostainer (Thermo Fisher Scientific, Inc). Sections were sealed and covered using the automated Thermo Clearview Coverslipper (Thermo Fisher Scientific, Inc).

Expression levels were assessed according to the chromagen signal using a simple immunoreactivity scoring system: 0 no expression, 1 (25-50% cells positive), 2 (50-75% cells positive) and 3 (75-100% cells positive).

2.7.2 Immunofluorescence

Immunofluorescence was used for semi quantitative analysis of Hh pathway expression in SGC in comparison to nodBCC. FFPE sections were cut (5 μm) and antigen retrieval carried out as described above. Blocking of the sections occurred with 5% goat serum for one hour followed by primary antibody staining as mentioned. Secondary antibody staining utilised AlexaFluor-568 (Invitrogen™ by Life technologies™). Nuclei were counterstained with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; Sigma-Aldrich®) at 1:1000 dilution to highlight the number of cells present and aid accurate semi-quantification. Coverslips were placed using VECTASHIELD® mounting medium (Vector Laboratories). Negative controls omitted the primary antibody and were essential for removal of background fluorescent noise to allow for accurate signal quantification and prevention of false positive acquisition.

Confocal Microscopy and fluorescence signal quantification

Immunofluorescence sections were examined using the Zeiss LSM710 Meta confocal laser microscope (Carl Zeiss Microscopy GmbH) and a ZEN configuration tool (Carl Zeiss Microscopy GmbH) which is a digital image processing software that produced an output TIFF image. Negative control slides were used to remove background fluorescence and set up the acquisition conditions that subsequently remained constant throughout the imaging process. Images were taken at 200 X magnification. Output TIFF images were analysed using ImageJ (http//imagej.nih.gov/ij) to quantify antibody expression. For semi-quantification of antibody expression, fluorescence intensity was determined in regions of interest (ROIs), ensuring a standardised area size whilst containing the same number of nuclei as determined by the DAPI staining. Furthermore, 3 separate ROI were taken for each area in each sample (such as 3 tumour ROI and 3 stroma ROI in nodBCC or SGC samples) to obtain a mean signal for comparison.

Statistical analysis

Immunofluorescence data are expressed as mean± standard error of the mean (SEM) of a given number of observations. Comparison between groups was made using the Student

t-test. A p-value of less than 0.05 was considered to be significant. Figures show the mean with error bars representing SEM.

2.7.3 Protein extraction from cell models and Western blotting

Protein extraction

Cells were harvested at ≥ 70% confluence using 50 mM sodium hydroxide at 100°C. Following a wash with phosphate buffered solution (PBS), 220 µl of heated protein lysis buffer (50 mM Tris-HCL, 2% sodium dodecyl sulphate (SDS) at pH 8.0) were added to the cells ensuring the cell surface was covered in its entirety before cells were removed with the cell scraper. Removed cells were placed in a 1.5 µl microcentrifuge tube and heated to 95° for 2 minutes. Passing the lysate through a needle and syringe reduced viscosity and 20 µl of lysate was removed for quantification using the Bio-RadDC Protein Assay detection kit (Bio-Rad). A mixture of 20 µl protein, 98 µl solution A (alkaline copper tartrate) and 2 µl solution S (SDS) was developed for 3 minutes before 800 µl of solution B (Folin) was added. Any protein binds to copper tartrate which in turn reduces Folin thereby turning the solution blue. The intensity of the blue correlates with the amount of protein present. The blue mixture was transferred to a 1 ml clear plastic cuvette and the protein absorbance (AU) measured using a spectrophotometer at a wavelength of 655 nm. A standard curve was obtained from known concentrations of bovine serum albumin provided and the protein concentration calculated in µg/µl. Samples were then stored at -80°C.

Western blot

Pre-cast Any kD™SDS-polyacrylamide gel (Bio-Rad Laboratories, Inc) was used to separate proteins. SDS interacts with protein to enable it to have a negative charge that is attracted to the cathode of the electric current. Twenty micrograms of protein was heated for 5 minutes at 95°C and mixed with 5X loading dye (100 mM Tris-GCL pH 6.8, 4% SDS. 200 mM dithiothreitol, 0.2% bromophenol blue and 20% glycerol). Exact volume varied in order to get 20 µg of protein per lane which was quantified using the Bio-RadDC Protein Assay detection kit (Bio-Rad Laboratories, Inc). DualColor Prestained Protein Ladder (Bio-Rad Laboratories, Inc) was used for the identification of protein size. The gel was placed into a mini-PROTEAN tetra cell system (Bio-Rad Laboratories, Inc) and run for 1 hour and 30 minutes at 60 volts in 1x Nupage MES SDS

running buffer (Life technologies™). Protein was transferred onto an Immun-Blot™ PVDF membrane (Bio-Rad Laboratories, Inc) using a Trans-Blot SD semi-dry transfer cell (Bio-Rad Laboratories, Inc) and Nupage Transfer Buffer (Life technologies™) at ~8V for 1 hour . Subsequently, the membrane was blocked overnight at 4ºC in blocking solution (0.4% PBS/Tween (PBST), 5% dry milk). The following morning, the membrane was washed five times in PBST and then incubated for 2 hours at room temperature with the antibody of interest (1:500 anti-PTCH, 1:200 anti-SMO, 1:1000 anti-Gli1 and 1:500 anti-Gli2). The membrane was then washed five times in PBST for 5 minutes, and incubated with 1:10 000 conjugated secondary anti-rabbit IgG horseradish peroxidise (HRP) (Dako) for 1 hour at room temperature. A further five washes of PBST for 5 minutes before the use of a ECL™ Prime western blotting detection system kit (GE Healthcare) which produces a chemiluminescent signal on binding to HRP. Signal was detected using the ChemiDoc MP Imaging system (Bio-Rad Laboratories, Inc) and ImageLab software (Bio-Rad Laboratories, Inc). Membranes were stripped of primary antibody using Restore™ Western Blot Stripping Buffer (Life technologies™) at 55ºC for 30 minutes, washed five times for 5 minutes in PBST and reprobed with 1:5000 polyclonal anti-β-actin antibody (Sigma-Aldrich®) as a loading control for each sample.

2.8 Cancer cell lines

2.8.1 LNCaP-Gli1 cancer cell line

This cancer cell line that ectopically expresses Gli1 was previously created in our lab by Dr Sandeep Nadendla and details on the methodology and validation are discussed elsewhere.(Nadendla et al., 2011) In summary, amphotropic viral particles containing pBABE-PURO-Gli1 were generated using the Phoenix packaging cell line (Nolan laboratory, Stanford, USA). Parental LNCaP cells (obtained from Sigma-Aldrich®) were exposed to viral particles with polybrene and centrifuged at 300g at 32°C for 1 hour. An elapse of 72 hours occurred prior to selection using puromycin to allow for the control, non-transduced cells to expire.

2.10 General solutions and media

Methyl green

0.1M Sodium Acetate Buffer, pH4.2

1.36g Sodium acetate, trihydrate (MW 136.1)

100ml dH_20

Mix to dissolve and adjust pH to 4.2 using glacial acetic acid

Methyl Green Solution (0.5%)

0.5g methyl green (ethyl violet from Sigma)

Dissolve in 100ml 0.1M sodium acetate buffer pH 4.2

PBS (pH 7.3)

137mM Nacl

2.7mM KCL

4.3 mM $Na_2HPO_4 \cdot 7H_20$

1.4 mM KH_2PO_4

Tris-HCL, 1M

121g Tris base in 800ml H_20

HCL to adjust pH

d H_20 to make 1L

Agarose Gel DNA runs(large)

300mls x1 TAE buffer

4.5grams Agarose

15ul gel red

50X TAE

242g Tris base

57.1ml glacial acetic acid

37.2g $Na_2EDTA \cdot 2H_20$

dH_20 to 1L

Antibody	Concentration	Ag retrieval
PTCH1	1 in 1500	High pH 9, 20mins 97°C
SMO Ab72130	1 in 200	High pH 9, 20mins 97°C
Gli1 Ab134906	1 in 250	High pH 9, 20mins 97°C
Gli2 Ab26056	1 in 150	Low pH 6, 20 mins 97°C
Alexa Fluro 568 A11011	1 in 500	------
GPC3 Ab66596	1 in 200	Low pH 6, 20 mins 97°C
HIST1H2BM Ab1790	1 in 1000	Low pH 6, 20 mins 97°C

CHAPTER THREE

GENETIC VARIANTS IN BASAL CELL CARCINOMA (BCC)

3.1 Aims of the BCC study

The main aim of the study was to identify the genetic makeup of sporadic, periocular BCC, specifically in reference to the different histological subtypes that confer a dissimilar prognosis. Although malignant, the nodular subtype is an indolent tumour that is readily amenable to curative surgical excision, however, the defect often requires reconstruction. In contrast, the morphoeic subtype is more locally aggressive, requires more mutilating surgery and in the worst case, requires the removal of a normal seeing eye (exenteration) to save life. Deciphering the genetic makeup of the more aggressive subtype by comparing it to its more benign counterpart was the initial objective. BCC from the so-called H zone on the face are classed as high risk, possible due to the delicate local structures preventing wide excision margins or due to their genetic makeup being different. Up to a third of tumours arise in non-UV exposed areas and so BCC elsewhere (elseBCC; defined as non-periocular BCC with intermittent exposure to UV light) may be genetically different and speculated to show a reduced UV light induced mutational signature and overall mutation burden. Thus, nodBCC from non-periocular sites were included in the study to see if they differ from those in the periocular region. It is hypothesised that H zone periocular tumours will have a higher UV burden than elseBCC and that the morphoeic subtype will have even more than nodBCC taken from any location. In addition, histological samples from the same individual can demonstrate both nodular and morphoeic subtypes and it is possible that morphoeic in this case is a subsequent development or evolved branching of the nodular subtype. Thus, a greater mutational burden with a more differential expressed transcriptome is expected. A discovery approach (Figure 3.1) using whole exome sequencing of 10 periocular mBCC (including 1 stroma) and 10 nodular BCC comprising of 5 periocular nodBCC and 5 non-periocular nodBCC (elseBCC) was carried out to address these issues. WES findings are summarised in this chapter. After looking at candidate genes, further examination using RNA sequencing (Chapter Four) of 3 periocular nodular and 3 morphoeic subtypes was performed to corroborate the exome sequencing data. The surrounding normal stromal

tissue was used as the reference, but also to see if the morphoeic stroma milieu possessed any changes that may explain its locally aggressive infiltrative nature by comparing it to the nodular stroma. On light microscopy, it has been long suspected that the surrounding stroma of the morphoeic subtype was different as a strong inflammatory reaction containing fibroblasts is seen (Figure 1.6). In light of the nucleic acid analysis, supportive evidence from protein expression (Chapter Four) was sought, especially in the hedgehog signalling pathway for both the tumour and stroma of human archival samples.

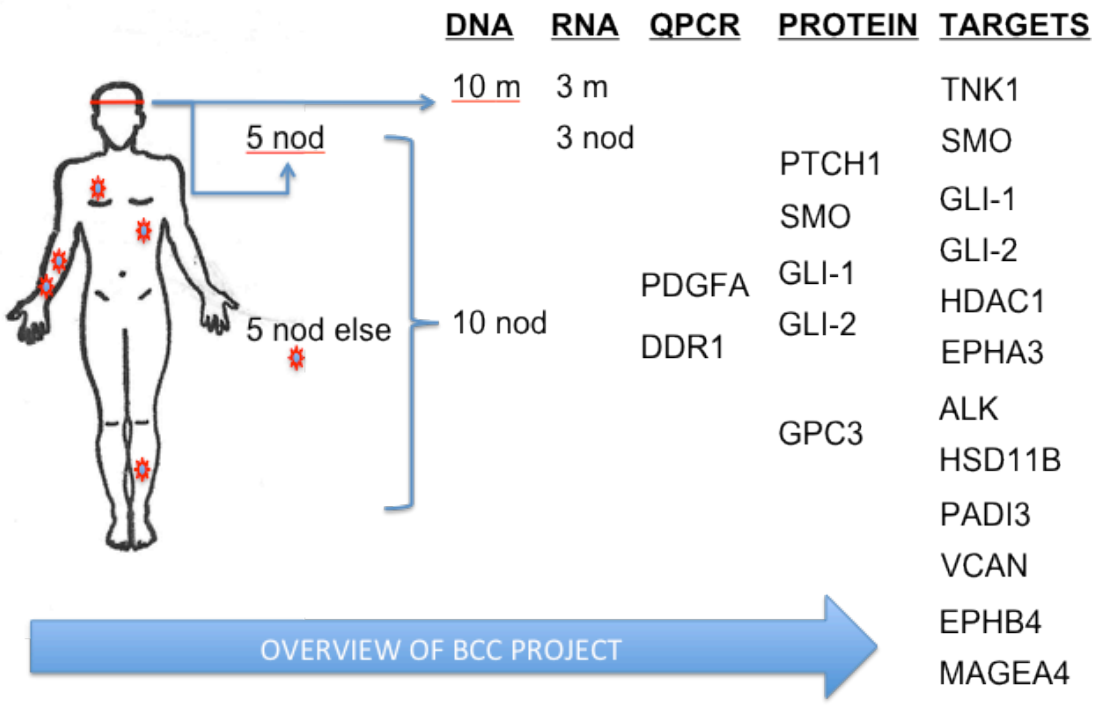


Figure 3.1 BCC project. Overview of the BCC project highlighting the stepwise progression starting with genetic variant detection in the DNA, followed by genetic expression using RNAseq and validation using quantitative PCR plus protein expression. Resultant treatment targets are summarised at the end of chapter 4. Nod, periocular nodular BCC; nod else, nodular BCC not within the high-risk H zone, m, periocular morphoeic BCC.

RESULTS

3.2 Whole exome sequencing in basal cell carcinoma (BCC)

Whole exome sequencing (WES) has massively increased the wealth of information on the mutational burden of cancer, but it can also add to the confusion. Large collaborations, for example the cancer gene atlas (TCGA), have got together to pool data in the hope that the larger numbers would allow the driver mutations to become more prominent with the background passenger mutations fading away. However, the large accumulation of sequencing data has not increased the sensitivity (number of driver mutations) and specificity (from passenger mutations) as desired. Despite correction for length and low false discovery rates being applied, very large genes such as *TTN* and *MUC16*, genes spanning long introns (for example *CSMD*) and those with unexplained biological significance like the olfactory receptors consistently come out on top.(Lawrence et al., 2013) Taking this awareness into account, the WES data is presented in a number of formats and different databases were consulted to look for potential driver mutations. Therefore, simple data such as frequency noted in the tumour is still shown and genes with odd biological function included. Genes of interest are described with a paragraph to introduce their known function and their relation to cancer given. Pathway grouping was another method of linking biologically similar mutations, however, tumours can act through non-canonical pathways (for example *KRAS* in pancreatic cancer). Cancer is heterogeneous and ultimately each patient is likely to be given personalised treatment according to their own mutational burden and expression in the future.

3.2.1 Clinical and histological features of BCC patients

The clinical and histological features of the BCC patient cohort used in the study are shown in Table 3.1 and Figure 3.2 respectively.

Sample	Age	Gender	Location	Sample	Age	Gender	Location
M1	85	M	RLL	**Nod1**	70	M	LLL
M2	47	F	LLL	**Nod2**	80	M	RLL
M3	78	F	LLL	**Nod3**	71	M	RMC
M4	71	M	LMC	**Nod4**	72	F	LLL
M5	83	F	LMC	**Nod5**	46	F	RLL
M6	61	M	RLL	**Else1**	74	M	Chest
M7	71	F	RMC	**Else2**	89	M	Chest
M8	78	F	RUL	**Else3**	83	M	Forearm
M9	62	F	RUL	**Else4**	87	M	Forearm
M10	87	F	RMC	**Else5**	52	M	Shin

Table 3.1 Clinical features of BCC patients. Summary of the clinical features for each BCC histological subtype that underwent whole exome and RNA sequencing. M, morphoeic BCC, Nod, nodular BCC, Else, non-periocular nodular BCC found outside the H-zone; LLL, left lower eyelid; RLL, right lower eyelid, LUL, left upper eyelid; RUL right upper eyelid.

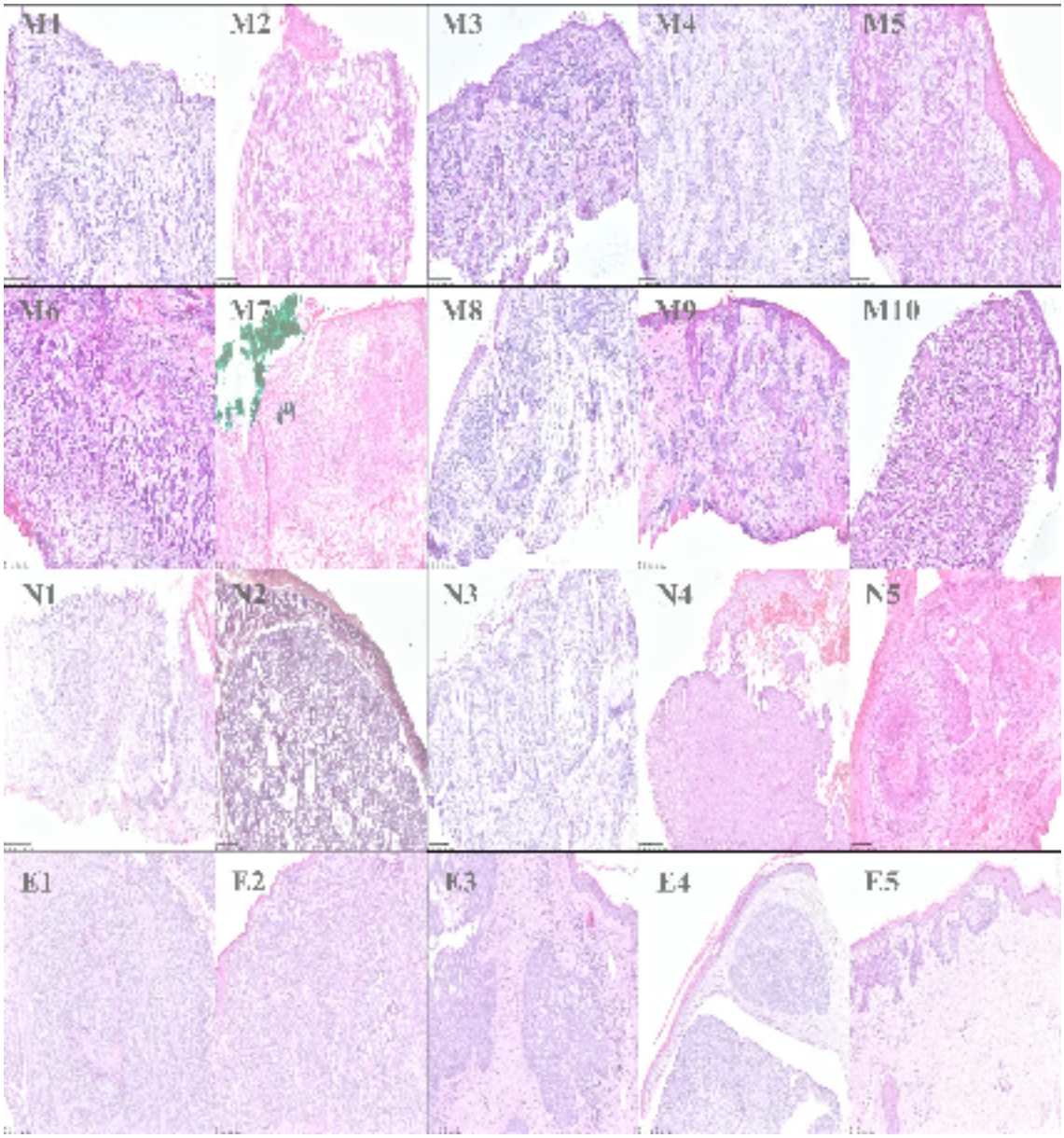


Figure 3.2 BCC patients recruited into the study. Histological collage of BCC samples presented as Haematoxylin and Eosin stained sections (H&E). All samples were initially processed as fresh tissue with fast H&E staining to identify tumour for laser capture. A few cases went on to FFPE processing with further H&E, hence the difference in quality between the samples. M, morphoeic BCC, N, nodular BCC, E, nodular BCC not within the H-zone. Scale bar at 100um.

3.2.2 Read level quality control metrics

Whole exome sequencing was performed on the Illumina HiSeq2000 platform. Read files (FASTq) were generated from the sequencing platform and read level quality control metrics provided by FastQC, a Babraham bioinformatics quality control tool for high throughput sequencing data (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw data from the platform is summarised in Figure 3.3. Quality control metrics are presented including mean read base quality (Phred score) (Figure 3.4), per sequence quality (Figure 3.5) and GC content per base (Figure 3.6). The Phred score is a measure of the quality of the identification of the base by the Illumina platform. A score of 30 indicates a 1 in 1000 probability of an incorrect base call conferring a base call accuracy 99.9%. Throughout the read the Phred score is approximately 32. The per sequence quality represents the quality of each read with the count representing the number of sequences. As for the individual base, a mean sequence quality (Phred) of greater than 30 indicates a good average quality per read. In addition, the graph highlights any subset of the sequences with that have poor quality by having a small peak below 20 which is not demonstrated on these samples. The per base GC content should reflect the organism's genome under investigation and is between 42-48 per cent in humans.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Number of Variations	70,432	66,407	69,223	73,527	71,078	69,867	70,061	70,694	69,589	70,504
Variations Overlapping Genes	70,010	66,035	68,800	73,057	70,655	69,455	69,678	70,289	69,185	70,054
Variations Overlapping Transcripts	70,010	66,035	68,800	73,057	70,655	69,455	69,678	70,289	69,185	70,054
Variations Overlapping Regulatory Regions	16,587	15,437	15,899	17,158	16,721	16,224	16,662	16,412	16,249	16,398
Variations Overlapping Protein Domains	66,468	62,803	65,546	69,439	67,240	65,991	66,238	66,751	65,704	66,594
Intergenic Variations	422	372	423	470	423	412	383	405	404	450
Variations With Predicted Serious Consequences	13,532	13,733	14,863	13,516	13,573	13,218	13,446	13,280	13,377	13,419
Variations With Other Predicted Consequences	56,478	52,302	53,937	59,541	57,082	56,237	56,232	57,009	55,808	56,635
Non-Synonymous Coding Variations with Consequences	12,799	12,998	14,031	12,775	12,790	12,488	12,720	12,548	12,678	12,703

Figure 3.3 All variants in mBCC. Metrics of all variants identified in WES including those from the single nucleotide polymorphism database (dbSNP) release 138.

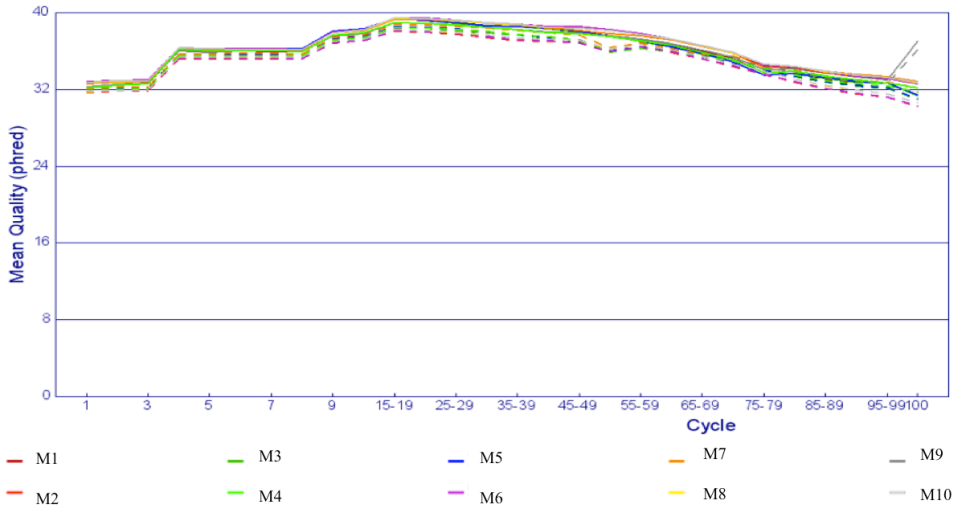


Figure 3.4 Mean Phred score for mBCC. Plot of the mean read base quality (Phred score) at each position of a read (cycle) for all ten mBCC samples. Mean quality decreases towards the end of the read as expected.

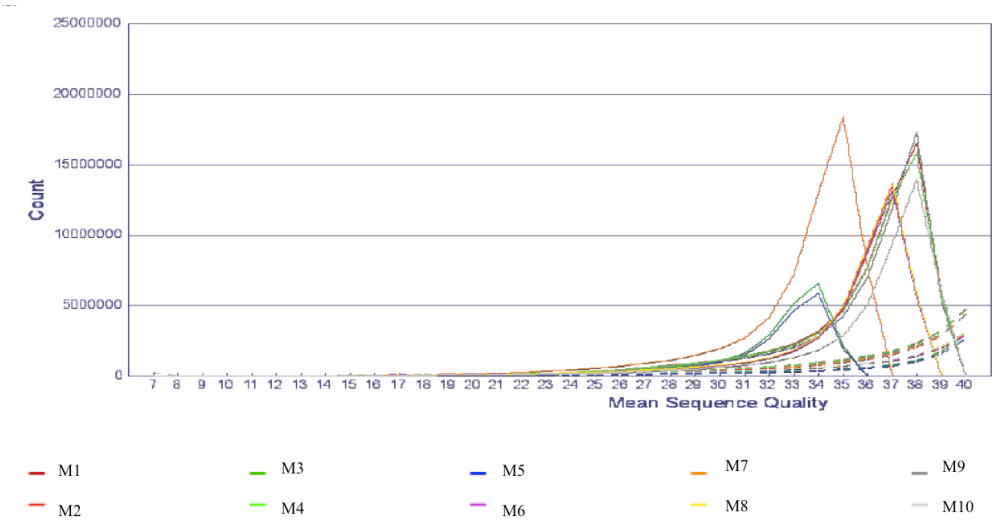


Figure 3.5 Per sequence quality in mBCC. Counts of the mean sequence quality (Phred score) for all individual samples' reads. If mean sequence quality is greater than 30 it indicates good quality and is for all samples.

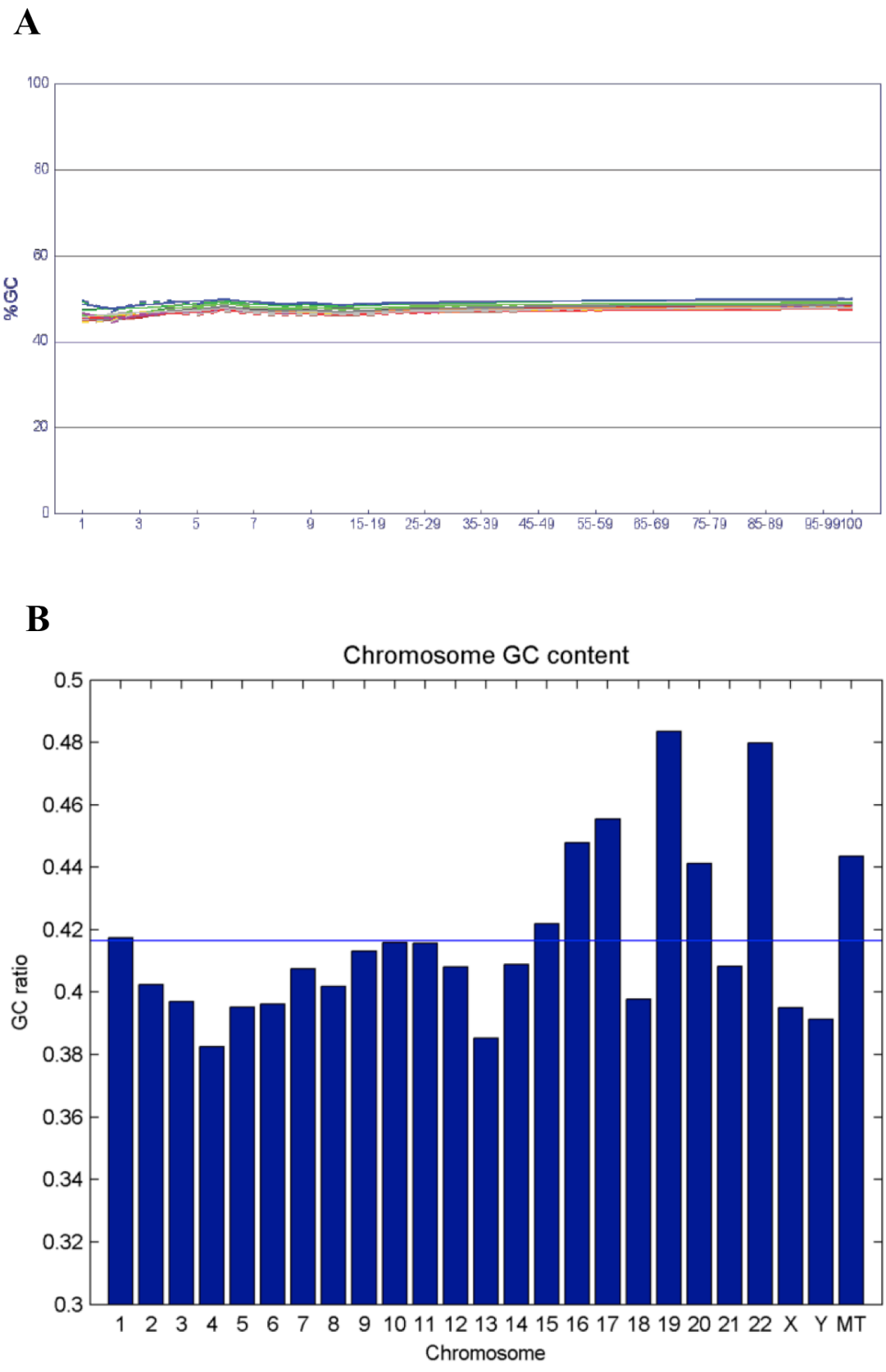


Figure 3.6 Per base GC content for mBCC. (A) Plot of percentage GC content at each position of a read (cycle) for all ten mBCC samples. It reflects the GC content for humans (B) between 42-48 percent. Figure 3.6B adapted from Romiguier et al. (Romiguier et al., 2010)

3.2.3 Mutational profile of BCC

A general overview of the WES data was performed to see how many variants were detected in each of the BCC subtypes giving an overall mutational burden within the tumour. Subsequently, UV signature analysis was performed to see how many variants were UV induced and whether this varied across the different subtypes.

Mutational burden

The total number of somatic coding mutations filtered for at least x20 depth for all ten nodular BCC was 18,446 and for mBCC 9564 (Table 3.2). Nodular elsewhere had a total burden of 10,368 compared to periocular nodBCC total of 8078. Periocular mBCC displayed a marginally reduced mutational burden to nodBCC (P=0.085) refuting the hypothesis that the more aggressive tumour would contain a greater number of mutations. Interestingly, the burden of mutation burden on normal eyelid skin alone is just under 1000, which highlights that a large number of the changes detected within these tumours are passenger mutations that are gathered in the skin. (Martincorena et al., 2015) Thus, it makes it harder to decipher the meaningful driver mutations.

Sample	Total	Non-syn	Syn	Sample	Total	Non-syn	Syn
M1	3107	1817	1018	**Nod1**	1936	1165	771
M2	958	560	327	**Nod2**	609	363	185
M3	1025	603	325	**Nod3**	2393	1402	991
M4	445	249	137	**Nod4**	1378	789	589
M5	611	375	180	**Nod5**	1762	1059	703
M6	166	93	53	**Else1**	1728	1017	527
M7	935	556	284	**Else2**	605	368	161
M8	89	58	28	**Else3**	213	133	50
M9	1132	702	340	**Else4**	4123	2348	1321
M10	1096	615	378	**Else5**	3699	2196	1124

Table 3.2 Mutational burden in BCC. Overall mutational burden, non-synonymous single nucleotide and synonymous single nucleotide variants for each sample that underwent WES after filtering for a minimum of x20 depth. M, morphoeic BCC, Nod, nodular BCC, Else, nodular BCC no in the H-zone; non-syn, non-synonymous single nucleotide variant; syn, synonymous single nucleotide variant.

UV burden and signature

Ultraviolet light can be genotoxic to DNA, producing cyclobutane pyrimidine dimers (CPD) and pyrimidine (6-4) pyrimidone photoproducts (64PPs) at dipyrimidine sites.(Ikehata and Ono, 2011) A specific type of base substitution occurs, a cytosine to thymine (C>T) at dipyrimidine sites. This especially occurs when the cytosine is preceded by another pyrimidine (in a CpC or TpC context) as this makes it easier for the dimer production to occur, as the dipyrimidine is not just nearby, but adjacent (CX or XC, where X= C or T). Moreover, these changes occur more frequently on the non-transcribed strand, as there is preferential nucleotide excision repair on the transcribed DNA strand (transcription coupled repair). Another, less common change is a tandem base substitution where two neighbouring cytosine (dinucleotide) are replaced by thymine (CC>TT). Methylated cytosine's at CpG islands seem to be more susceptible to CPD formation, especially from UVA and UVB light. The UV signature of normal skin is 55% and the majority tend to be a dinucleotide change (CC>TT).(Martincorena et al., 2015) The UV signature across all subtypes showed a similar change at 85% for mBCC, 88% for nodBCC and 85% for elseBCC (Figures 3.7 and 3.8). This refutes the hypothesis that the aggressive nature of mBCC is related to both UV burden and locality. It was thought that those on the face (within the high-risk H-zone) would be more exposed to sunlight over the course of time than those that are covered most of the time, hence would show an increase in UV signature along with tumour total mutation burden. mBCC 10 was an outlier and only demonstrated a UV signature percentage of 45% reflecting normal skin burden percentage, but a greater number of potentially pathological changes compared to normal skin. The lack of UV signature may represent an early stage in the cancers development, the UV induced genes may not beneficial to the cancer and therefore not positively selected.

In addition, the mBCC histological subtype is a proven risk factor for recurrence and being a UV associated tumour, it was thought the aggressive nature would be related to more UV induced mutations, but this was not the case. (Mosterd et al., 2008) Nevertheless, it was postulated by Martincorena *et al* that every sun exposed skin cell acquires a new mutation in its genome for everyday of its life.(Martincorena et al., 2015) Moreover, it has been shown that as little as three driver mutations are required for the onset of lung and colorectal cancer.(Tomasetti et al., 2015) Thus, the normal skin mutational burden makes it more difficult to find the initiation drivers of BCC.

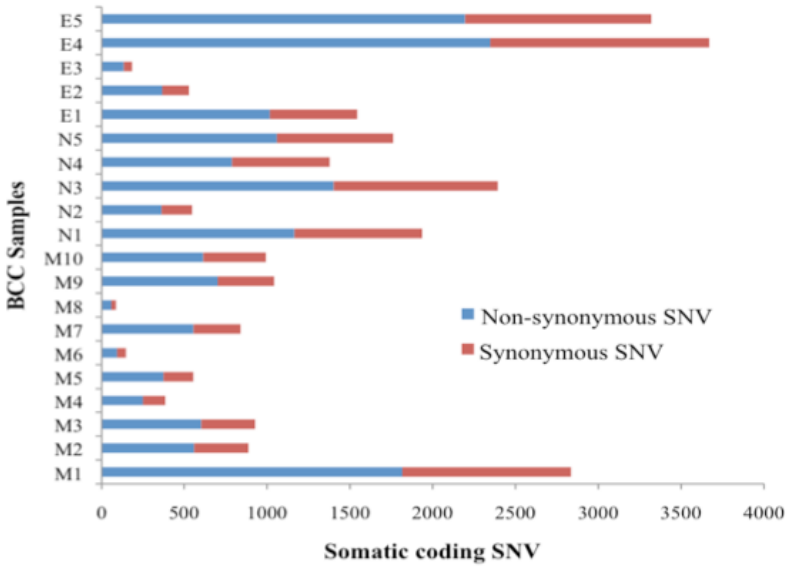


Figure 3.7 Mutational burden of basal cell carcinoma. Somatic coding single nucleotide variant (SNV) burden for each BCC WES sample with the blue portion of the bar representing non-synonymous SNVs and the red synonymous SNVs

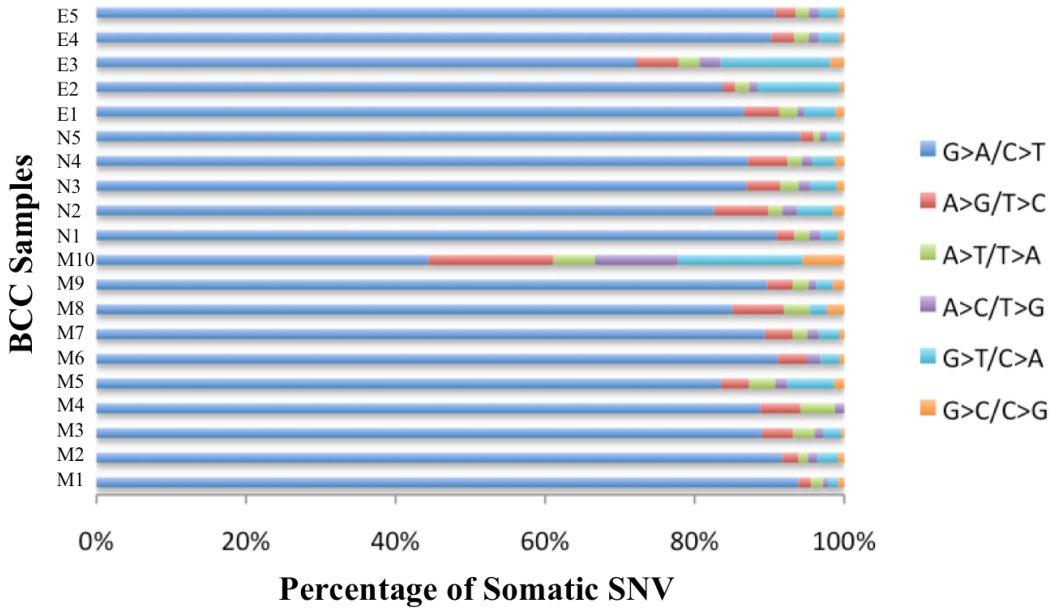


Figure 3.8 Pathological mutations and signature in BCC. Somatic single nucleotide variants are shown for each BCC sample with the percentage of each possible transition/transversion shown. M, morphoeic BCC; N, nodular BCC; E, non-periocular nodular BCC located elsewhere. Key: Pyrimidine transition, purine transition and four transversions.

3.2.4 Driver mutations in periocular morphoeic BCC

The ten mBCC samples were analysed together using two different mutational significance algorithms, namely MutSigCV and Intogen, resulting in 38 (Table 3.4, 3.5 and 3.6) and 17 (Table 3.7, 3.8) significant genes respectively. The striking result is that they shared just one gene, namely PTCH1, albeit the most reassuring one that confirms our samples were genuine BCC (Figure 3.9). Genes from both lists are explored in more detail in relation to their biological function and cancer association. Genes are then collated to identify the mutations in relation to their involvement in a molecular pathway (section 3.2.5).

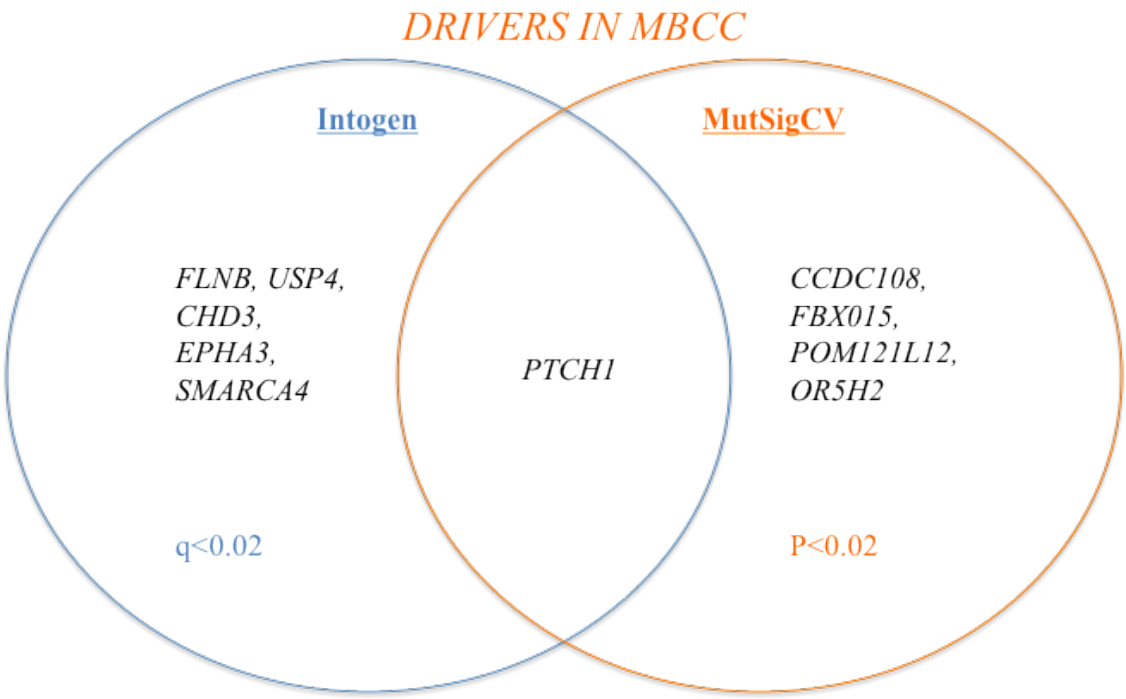


Figure 3.9 Driver genes in mBCC. Venn diagram highlighting the significant driver genes in mBCC using two algorithms, MutSigCV and Intogen, with shared genes within the cross over region**.** Q, false discovery rate; p, p-value.

3.2.4.i Shared driver mutation using MutSigCV and Intogen

Only one gene, *PTCH1* was identified using both algorithms (Figure 3.9). This may reflect the heterogeneity of the tumour whereby a multitude of different genes can lead to a morphoeic phenotype. Table 3.3 summarises the *PTCH1* variants that were found in the mBCC cohort with the predicted outcome for the variant using SIFT and Polypen.

Sample	Genomic / DNA level	Protein level	SIFT	Poly	Cos id
M1	c.988T>C	p.Thr230Ala	0.7, T	0.99, PD	
M2	g.98244414A>T (Intronic)	p.? (Splice distance 2)			
M3	c.1885T>A	p.Lys529*			144245
M4	c.3077C>G	p.Trp926Ser	0, D,H	1, PD	
M5	c.3419_3427delinsA	p.(Ser1040fs)			
M6	c.3059_3060insAGAAAGC	p.(Tyr920fs)			
M7	c.3549_3550insCA	p.(Val1083fs)			
M8	c.2536T>A	p.Lys746*			

Table 3.3 *PTCH1* mutations seen in mBCC cohort. The mutation DNA location and protein change are highlighted. * Stop/gain; fs, non-synonymous frameshift:stop/gain. Analysis from SIFT (D=damaging, T=tolerated, L=low probability, H=high probability); Poly=Polypen (PD=probably damaging, B=benign) and Cos id=COSMIC identifier.

3.2.4.ii Drivers identified using MutSigCV algorithm

Four driver mutations were discovered using the MutSigCV algorithm with a $P<0.02$ (Figure 3.9). However, due to only one gene being shared across the Intogen and MutSigCV, the data is presented in an alternative format to highlight any other potential driver genes. The algorithm is weighted to non-silent variants, but also taking into account silent variants (Table 3.4). If genes were organised using non-silent changes then they would be ordered differently and C6 becomes significant (Table 3.5). However, it also takes into account the size of the gene and its function otherwise TTN would come out as the most important gene (Table 3.6). Selected potential drivers are explored in more detail and listed below the tables.

Gene	Non-silent	Silent	P Value
PTCH1	8	1	1.61E-07
CCDC108	4	2	0.011
FBXO15	2	0	0.013
POM121L12	6	1	0.015
OR5H2	2	0	0.018
KRTAP12-3	1	0	0.02

Table 3.4 Drivers in mBCC using MutSigCV silent and non-silent mutations. Top six driver genes using MutSigCV taking into account silent and non-silent genes, and ordered according to P value with a $P \leq 0.02$.

Gene	Non-silent	Silent	P Value
PTCH1	8	1	0.000000161
C6	7	0	0.040
POM121L12	6	1	0.013
CCDC108	4	2	0.015

Table 3.5 Drivers in mBCC using MutSigCV non-silent mutations. Top six driver genes using MutSigCV taking into account non-silent frequency only, and ordered according to P value with a $P<0.02$.

Gene	Non-silent	Silent	P value	Gene	Non-silent	Silent	P value
TTN	52	15	9.99E-01	GPR112	7	1	2.80E-01
MUC16	17	4	1.00E+00	TEX15	7	1	3.69E-01
CSMD3	13	2	8.72E-01	RYR1	7	0	5.10E-01
SYNE1	10	5	9.61E-01	RP1	7	1	7.62E-01
CSMD1	9	2	9.99E-01	LRP1B	7	2	9.23E-01
PCDH15	9	5	1.00E+00	DNAH8	7	2	9.70E-01
PTCH1	8	1	1.61E-07	GPR98	7	2	1.00E+00
ANK2	8	2	4.96E-01	POM121L12	6	1	1.46E-02
BAI3	8	1	9.92E-01	PKD1L2	6	1	1.69E-01
C6	7	0	4.02E-02	COL11A1	6	0	3.07E-01
EPHA3	7	0	1.05E-01				

Table 3.6 Most mutated mBCC genes using MutSigCV ignoring size and function. Top 20 genes arranged according to the number of non-silent gene mutations in MutSigCV ignoring size and function.

3.2.4.iii Drivers identified using Intogen algorithm

WES data was put through the Intogen algorithm too and the potential driver genes listed in the Table 3.7 using a false discovery rate (q-value) of < 0.05. TTN is included, as it is not removed in the same way as MutSigCV, however, although it contains many variants, it is not attributed driver status by Intogen program. A list of known drivers that were detected is listed in Table 3.8 but many are not significant. A more stringent q value was used for the Venn diagram (Figure 3.9) to narrow the number of potential drivers. These among other selected genes are discussed in more detail below the tables.

Gene	Freq	Driver	q-value	Gene	Freq	Driver	q-value
PTCH1	8	HCD	1.06E-10	NFS1	3		0.038
FLNB	5		1.51E-8	EFTUD2	2	CD	0.038
TTN	8		4.44E-8	HECTD4	4		0.038
USP4	4		7.71E-4	SMARCA4	4	HCD	0.038
CHD3	2	CD	1.39E-3	FNIP1	2		0.042
EPHA3	3	CD	0.007	CHKB	2		0.043

Table 3.7 Intogen derived drivers in mBCC. Top 12 driver genes arranged using Intogen and an fm-bias (q-value) value <0.05. CD, connected to a known driver; HCD, high confidence driver; KD, known driver, Freq, frequency.

Gene	Freq	q-value	Gene	Freq	q-value
PTCH1	8	<0.001	AHNAK	2	0.192
CHD3	2	0.001	MED12	2	0.346
EFTUD2	2	0.039	ARID4B	3	0.346
ARHGAP35	4	0.039	ATR	2	0.487
CREBBP	3	0.066	PLXNA1	2	0.487

Table 3.8 Known Intogen drivers found in mBCC. Top 10 known driver genes using Intogen arranged in order of reducing fm-bias (q-value) significance.

3.2.5 Morphoeic BCC molecular pathway analysis using WES data

Potentially pathological variants were collated according to pathway involvement to try and identify any associations. Intogen was used to collated the variants and list the most significant altered pathways according a q <0.05 (Table 3.9). In doing so, four pathways came out on top namely the hedgehog, BCC, pathways in cancer and Wnt signalling pathways (Table 3.9). The mutational burden of four biologically significant pathways were picked for further assessment to see which genes were changed and included Hh, NK cell, Wnt and FC epsilon RI signalling pathway (Tables 3.10 – 3.13). The later pathway was just non-significant at a q value of 0.06; however, it is a potential candidate pathway as it is involved in the release of inflammatory mediators and cytokines (Figure 3.9). Another potential pathway (axonal guidance) not picked using Intogen q<0.05, but has been associated with pancreatic cancer and is altered in our RNAseq analysis is mentioned below in more detail.

Identifier	Description	Q-value	Freq
hsa04340	Hedgehog signaling pathway	7.23E-10	9
hsa05217	Basal cell carcinoma	2.33E-08	9
hsa05200	Pathways in cancer	2.99E-05	10
hsa04310	Wnt signaling pathway	0.000536	9
hsa05032	Morphine addiction	0.0062	8
hsa05414	Dilated cardiomyopathy	0.0112	8
hsa04976	Bile secretion	0.0146	8
hsa05166	HTLV-I infection	0.0234	9
hsa04916	Melanogenesis	0.0234	8
hsa04650	Natural killer cell mediated cytotoxicity	0.0417	8
hsa04974	Protein digestion and absorption	0.0488	9

Table 3.9 Driver pathways in mBCC. Intogen pathway prediction highlighting altered pathways using WES data from 10 mBCC tumours and a q < 0.05. Freq, the number of mBCC samples that the pathway is significantly altered.

Gene	Freq	Driver	Gene	Freq	Driver
PTCH1	8	Driver, Rec	CSNK1E	1	
SMO	1	Driver, Dom	PRKACB		
Gli1	1		PRKX	1	
Gli2			WNT2		
Gli3	1		PRKACA	1	
STK36	2		SUFU	1	Driver, Rec
LRP2	2		WNT5B	1	
WNT8A			BMP5	1	
CREBBP	2	Driver	PTCH2	1	
HHIP			WNT10A	1	
SIN3A	1		BMP6	1	
BMP4	1		WNT3A	1	
CSNK1A1	1		SHH	1	

Table 3.10 Hedgehog pathway and related genes in mBCC. Mutational burden in 10 mBCC tumours with respect to the Hedgehog signalling pathway (hsa04340). Specific variants are highlighted as possible driver mutations according to the Intogen algorithm along with their either dominant or recessive effect. Freq, frequency; Rec, recessive; Dom, dominant.

Gene	Freq	Driver	IFNA16	2	
ARAF	1		IFNA2	1	
FCER1G	1		IFNA4	1	
FCGR3A	2		IFNA7	2	
FYN	1		PIK3CD	1	
HRAS	1	HCD, Dom	PIK3CG	1	HCD
IFNA10	2				

Table 3.11 Natural Killer pathway in mBCC. Mutational burden in 10 mBCC tumours with respect to the NK pathway (hsa04650). Specific variants are highlighted as possible driver mutations according to the Intogen algorithm along with their either dominant or recessive effect. Freq, frequency; Rec, recessive; Dom, dominant; HCD, high confidence driver.

Gene	Freq	Driver	Gene	Freq	Driver
CREBBP	2		ROCK2	1	
NFATC2	2		APC	1	HCD,Rec
NFAT5	2		WNT10A	1	
RHOA	2	HCD	PPP2R1B	2	
TBL1X	2		PPP3CA	1	
PPP3R1	2		TP53	2	HCD,Rec
MAPK9	1		SMAD4	1	HCD,Rec
LRP6	1		TCF7L2	1	HCD,Dom
PRKACA	1		PLCB3	4	
EP300	1	HCD, Rec	WNT3A	1	
CHD8	1	HCD	CTBP1	1	
NFATC4	2		DVL3	1	
CSNK2A1	1		PRICKLE2	1	
PORCN	1		SOX17	1	HCD
AXIN1	1		PRKCB	1	
FZD3	1		SMAD3	1	
SFRP4	2		RAC3	1	
DVL1	1		VANGL1	1	
MAPK8	1		RUVBL1	1	
WNT5B	1		PRKX	1	
FZD10	1		FZD9	2	
PPP2R5D	1		CSNK2B	2	
CSNK1A1	1		CSNK1E	1	

Table 3.12 Wnt signalling and related genes in mBCC. Mutational burden in 10 mBCC tumours with respect to the Wnt signalling pathway (hsa04310). Specific variants are highlighted as possible driver mutations according to the Intogen algorithm along with their either dominant or recessive effect. Freq, frequency; Rec, recessive; Dom, dominant, HCD, high confidence driver.

Gene	Freq	Driver	Gene	Freq	Driver
INPP5D	2		VAV1	1	
FYN	1		PIK3R1	1	HCD, Rec
GAB2	1		MS4A2	1	
MAP2K3	1		FCER1G	1	
MAPK9	1		PRKCB	1	
SOS2	1		RAC3	1	
PIK3R2	2		PIK3CD	1	
PIK3CG	1	HCD	HRAS	1	HCD, Dom
MAPK8	1		PLA2G6	1	
MAPK14	1		PLCG2	2	
AKT3	1		LAT	1	

Table 3.13 Fc epsilon RI signalling in mBCC. Mutational burden in 10 mBCC tumours with respect to the Fc epsilon RI signalling pathway (hsa04664s). Specific variants are highlighted as possible driver mutations according to the Intogen algorithm along with their either dominant or recessive effect. Freq, frequency; Rec, recessive; Dom, dominant.

As mentioned the Fc epsilon RI signalling pathway was only just non-significant, but looking at the Kegg pathway in terms of flow you can see that alteration of the known driver genes highlighted in orange could have a profound affect (Figure 3.10).

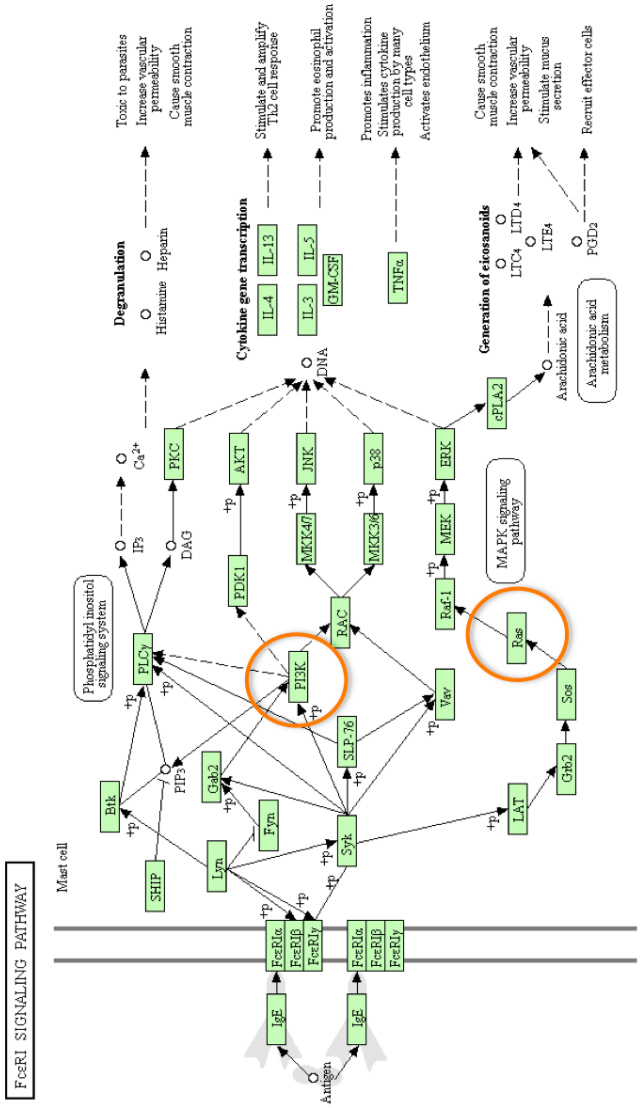


Figure 3.10 FC epsilon RI signalling pathway adapted from Kegg (hsa04664). Orange circles represent high confidence driver (HCD) genes identified using Intogen that were mutated in mBCC.

Axonal guidance pathway

Another possible novel pathway in mBCC is the axonal guidance pathway (Figure 3.11). Each gene is summarised in Figure 3.11 which visually summarises the pathway burden for each of the 10 mBCC samples. Although not within q <0.05 Intogen analysis, it is sufficiently affected to suspect potential involvement, especially as it is significantly modified at the RNA level too (see section 3.4.2). A Kegg pathway diagram (Figure 3.12) is also shown with orange circles highlighted the variant genes with the mBCC cohort.

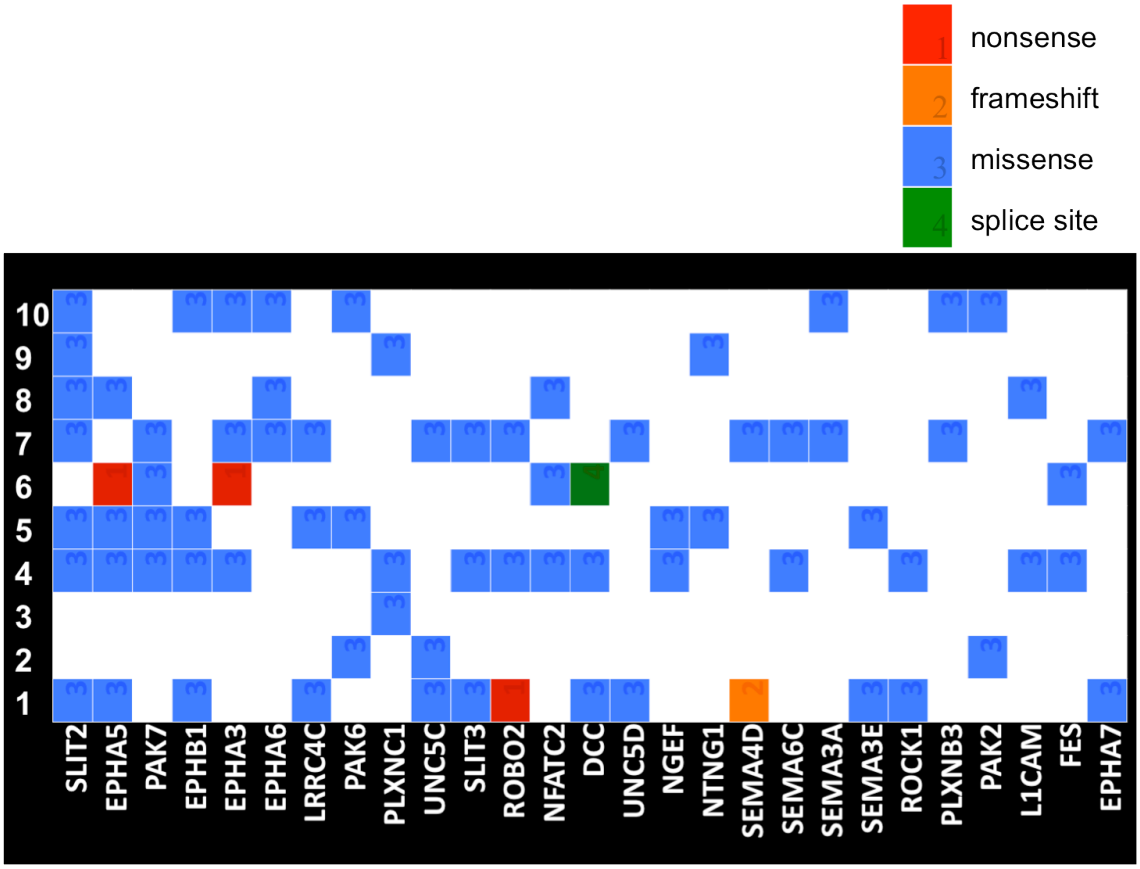


Figure 3.11 Axonal guidance in mBCC. Heatmap representing the axonal guidance pathway (genes along x-axis) mutations in all 10 mBCC (y-axis). Each colour represents a particular variant within the gene.

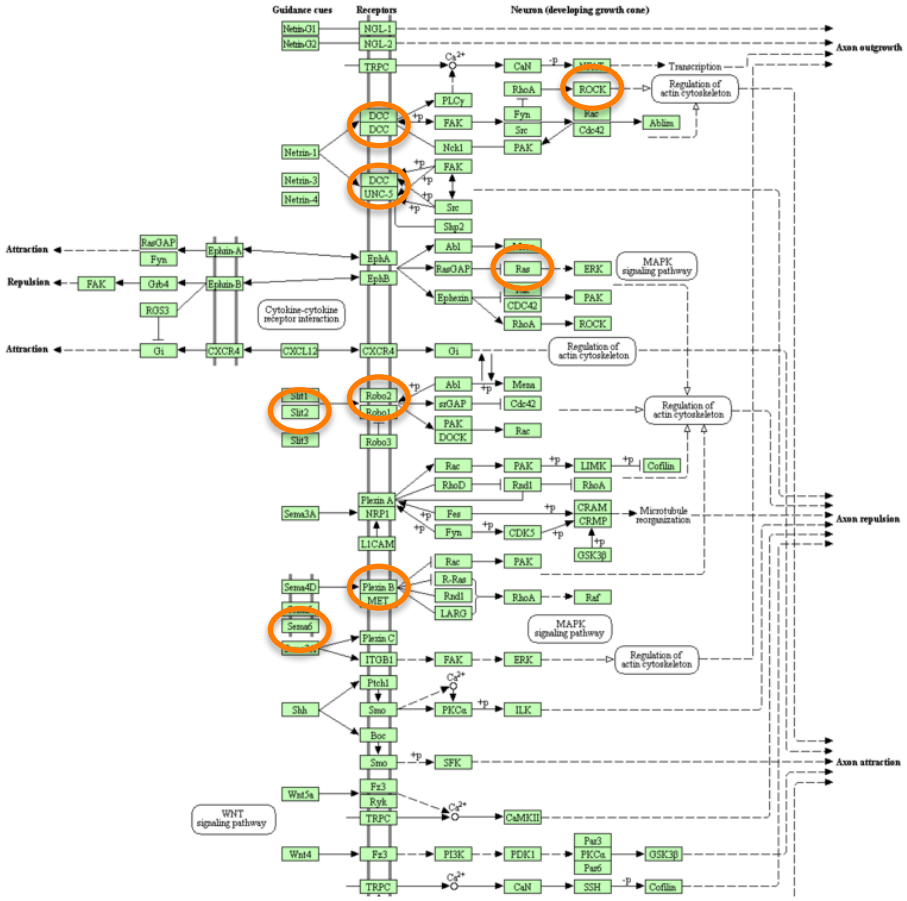


Figure 3.12 Axonal guidance pathway adapted from Kegg (hsa04664). Orange circles represents genes identified using Intogen that were mutated in mBCC. Also shown are the Hh and Wnt signalling pathways and how they all work together during axonal development.

3.2.6 Driver mutations in nodular BCC

Ten nodular BCC, 5 from the periocular region and 5 from elsewhere, were analysed together using two different mutational significance algorithms, MutSigCV and Intogen. This resulted in 78 and 176 (22 drivers) significant (P<0.05) genes using respectively. Restricting the significance cut-off to P<0.01 for the MutSigCV genes reduced the total to 8 genes (Table 3.14). Using a q-value (adjusted p-value taking into account the false discovery rate, FDR) of less than 0.05, the Intogen total reduced to 23 (7 drivers) top genes listed in Table 3.15. When comparing the two, the shared significant (P<0.05) mutations were found in six genes: PTCH1, TP53, FADS1, PPM1D, TAF4B, TRIP4 (Figure 3.13). Genes from both algorithms are explored in more detail in relation to their biological function and cancer association

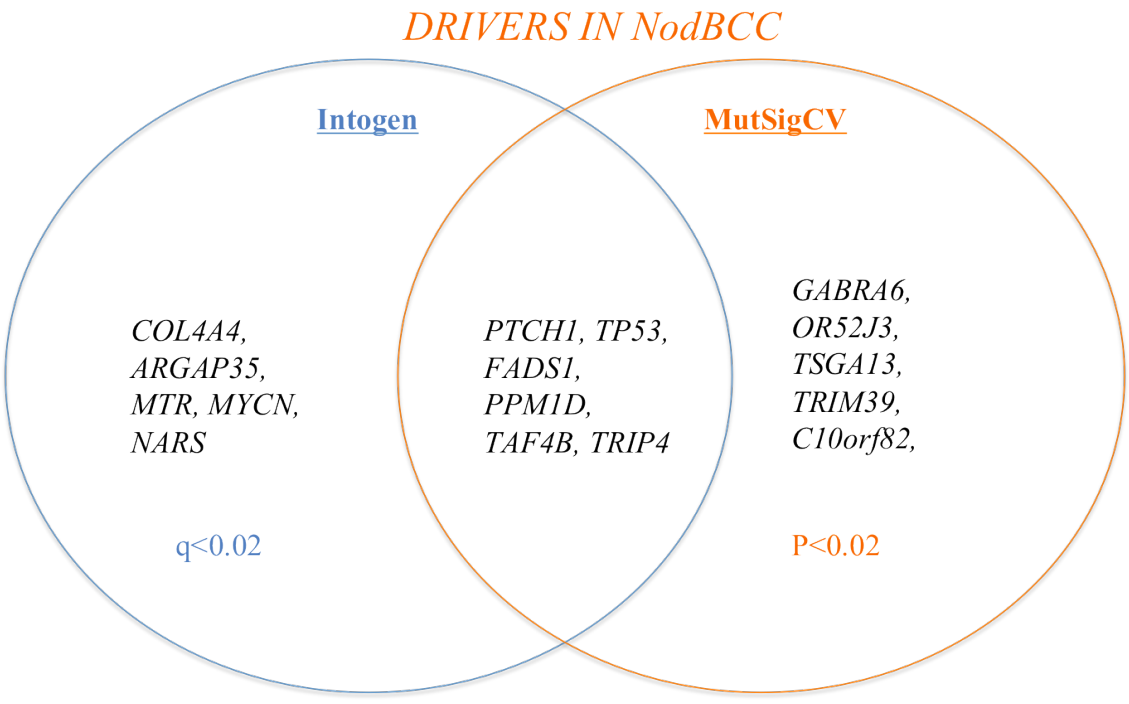


Figure 3.13 Drivers in nodBCC. Venn diagram highlighting the significant driver genes identified in nodBCC using two algorithms, MutSigCV (p<0.02) and Intogen (q<0.02) with shared genes in the cross over region.

3.2.6.i Shared driver mutations using MutSigCV and Intogen

Six genes were found to be potential drivers using both algorithms including PTCH1, TP53, FADS1PPM1D, TAF4B, TRIP4 and are discussed in more detail below (see 3.3.4).

3.2.6.ii Drivers identified using MutSigCV algorithm

Eleven genes were identified using MutSigCV and a P<0.02. The number of silent and non-silent variants are listed for the top 8 mutations in Table 3.14. Six of the genes shared with Intogen have been discussed and the five other genes are discussed in more detail below.

Gene	Non-silent	Silent	P Val	Gene	Non-silent	Silent	P Val
TP53	13	2		TSGA13	3	0	
PTCH1	9	1		TRIM39	3	0	
GABRA6	5	1		C10orf82	3	0	
OR52J3	4	0		FADS1	2	0	

Table 3.14 Drivers in nodBCC using MutSigCV silent and non-silent mutations. Top 8 potential driver mutations MutSigCV genes taking into account silent and non-silent genes ordered according to P value with a P<0.01

3.2.6.iii Drivers identified using Intogen algorithm

Twenty-three genes were identified using a q<0.05 and are summarised in Table 3.15. Several of the variants were also noted in databases such as COSMIC and an example of their identifying numbers is also mentioned. PTCH1 changes were replicated in 26 samples documented in COSMIC. Selected Intogen potential driver genes are listed below.

Gene	F	Dr	Database ID	Gene	F	Dr	Database ID
PTCH1	7	1	* (4) COSM96959	ITGAL	4	0	
			* (26) COSM10662	NHSL1	4	0	
TP53	7	1	* (2) rs28934573	SYNE1	8	0	rs201346604
COL4A4	3	0		RAB14	2	0	
TRIP4	4	0	rs144077712	IGSF10	3	0	
ARHGAP35	3	1		KANSL1	2	0	
MTR	3	0	rs142250261	PPM1D	2	1	
			*(8)COSM286164;	FADS1	3	0	rs142276858
TTN	9	0	* (5)rs145185269	TAF4B	2	0	
MYCN	4	1		ARID1A	2	1	COSM51453
NARS	2	0	ESP_18_55274765	IRF2	2	1	
PAM	4	0		TANC1	3	0	
ZEB1	4	0					

Table 3.15 Intogen drivers in nodBCC Top 23 genes sorted using Intogen using a fm-bias q value <0.05 *(total number) multiple known variants were identified in the following databases COSMIC, dbSNP and ESP. An example of the known variant is also shown.

3.2.7 Nodular BCC molecular pathway analysis using WES data

Intogen pathway prediction was performed on nodBCC WES data across the ten samples. Forty-two pathways are significantly altered using a $q<0.05$ and the ones with a q value less than 0.02 are summarised in Table 3.16. Specific pathway variants for Hh, TGF-beta and TP53 are shown in Tables 3.17, 3.18 and 3.19 respectively.

Identifier	Description	Q- value	Freq
hsa05217	Basal cell carcinoma	2.25E-09	9
hsa04340	Hedgehog signaling pathway	1.61E-06	9
hsa05203	Viral carcinogenesis	2.46E-05	10
hsa04110	Cell cycle	8.26E-05	10
hsa04350	TGF-beta signaling pathway	1.2E-04	9
hsa04115	p53 signaling pathway	1.2E-04	10
hsa05200	Pathways in cancer	4.0E-04	10
hsa05414	Dilated cardiomyopathy	1.0E-03	10
hsa05213	Endometrial cancer	1.2E-0.3	9
hsa04270	Vascular smooth muscle contraction	1.3E-0.3	9

Table 3.16 Driver pathway in nodBCC. Intogen pathway prediction highlighting altered pathways using WES data from 10 nodBCC tumours.

Gene	Freq	Driver	Gene	Freq	Mutation
PTCH1	8	Driver, Rec	CSNK1E	1	
SMO	1	Driver, Dom	PRKACB		
Gli1	1		PRKX	1	
Gli2			WNT2		
Gli3	1		PRKACA	1	
STK36	2		SUFU	1	Driver, Rec
LRP2	2		WNT5B	1	
WNT8A			BMP5	1	
CREBBP	2	Driver	PTCH2	1	
HHIP			WNT10A	1	
SIN3A	1		BMP6	1	
BMP4	1		WNT3A	1	
CSNK1A1	1		SHH	1	

Table 3.17 Hedgehog pathway in nodBCC. The mutational burden in 10 nodular BCC with respect to the Hedgehog signalling pathway is highlighted. Specific variants are highlighted as possible driver mutations according to the Intogen algorithm along with their either dominant or recessive effect..

Gene	Freq	Driver	Gene	Freq	Driver
ACVR1C	1		MAPK3	2	
ACVRL1	1		RBL1	1	
AMH	3		ROCK1	3	
AMHR2	1		SMAD3	1	
BMP4	1		SMAD4	2	
BMP6	3		TGFB1	1	
CHRD	1		TGFB2	1	
DCN	1				

Table 3.18 TGF-beta pathway in nodBCC. Highlighting the mutational burden in 10 nodular BCC with respect to TGF-beta pathway (hsa04350).

Gene	Freq	Driver	Gene	Freq	Driver
APAF1	3	D	CCND3	1	
ATM	2	D	CCNG2	1	
ATR	2	D	CDK4	1	D
BAI1	3		TP53	7	D
CASP8	3	D	ZMAT3	1	
CCNB3	2				

Table 3.19 TP53 pathway in nodBCC. Highlighting the mutational burden in 10 nodular BCC with respect to TP53 pathway.

3.2.8 Locality and subtype specific drivers

Next, we performed the randomisation analysis (see Methods section 2.6.1) to identify significantly mutated genes that were related to a specific subtype or locality (Figures 3.14-3.16). MutSigCV mBCC specific gene included *CCDC108*, and nodBCC specific included *GABRA6*, *OR52J3*, *TRIM39* and *ADAM29* (Figure 3.14). Intogen mBCC specific genes included *HECTD4* and *FLNB* and nodBCC specific were *ZEB1* and *TP53* (Figure 3.15). Locality nodBCC genes include *SIPA1L1* and *SYNE1* whereas elseBCC revealed *ANK2* and *HIVEP1* (Figure 3.16). *TP53* and *PTCH1* were shared across both locations.

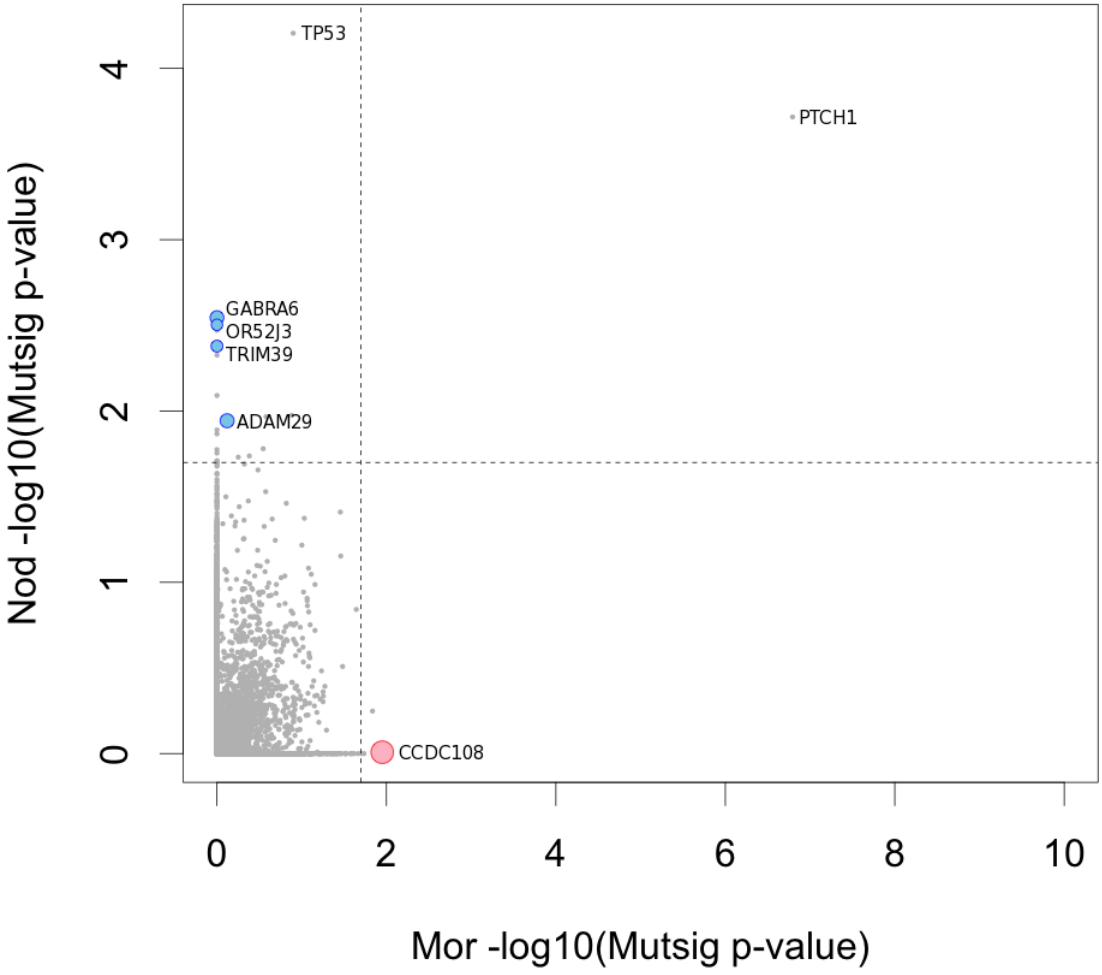


Figure 3.14 Randomised comparison of MutSigCV drivers in BCC. Drivers identified using MutSigCV were compared across the two subtypes, morphoeic and nodular BCC, and a randomised comparison test performed to discover those drivers that are more likely to be noteworthy than being randomly picked by chance during WES discovery. Genes close to an axis are specific for that subtype. Genes in the middle are shared between the two. The size of the circle indicates the confidence that the gene is a driver. Cut off P=0.02 (-log10(0.02)=1.7, the dotted line).

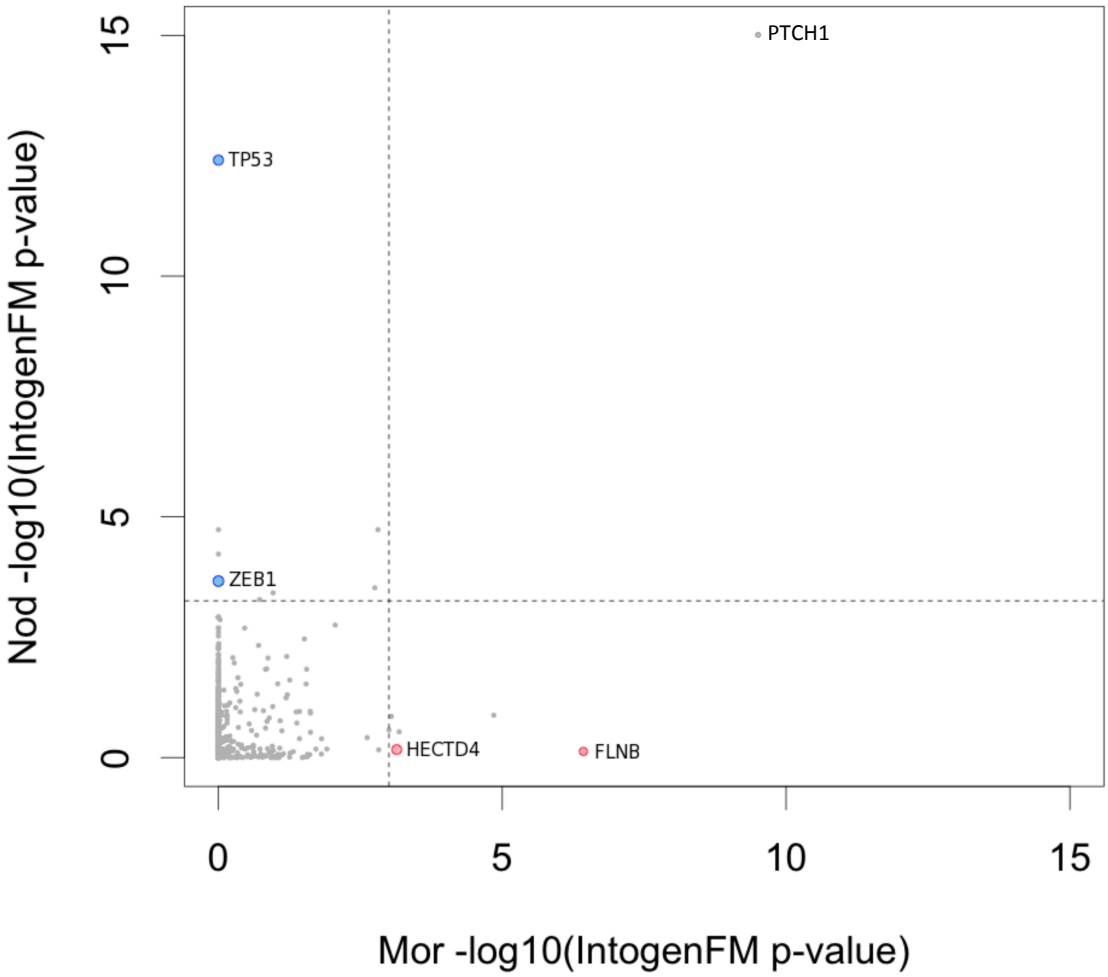


Figure 3.15 Randomised comparison of Intogen drivers to identify subtype specific genes in BCC. Drivers identified using Intogen were compared across the two subtypes, morphoeic and nodular BCC, and a randomised comparison test performed to discover those drivers that are more likely to be noteworthy than being randomly picked by chance during WES discovery. Genes close to an axis are specific for that subtype. Genes in the middle are shared between the two. The size of the circle indicates the confidence that the gene is a driver. Cut off P=0.001 (dotted line)

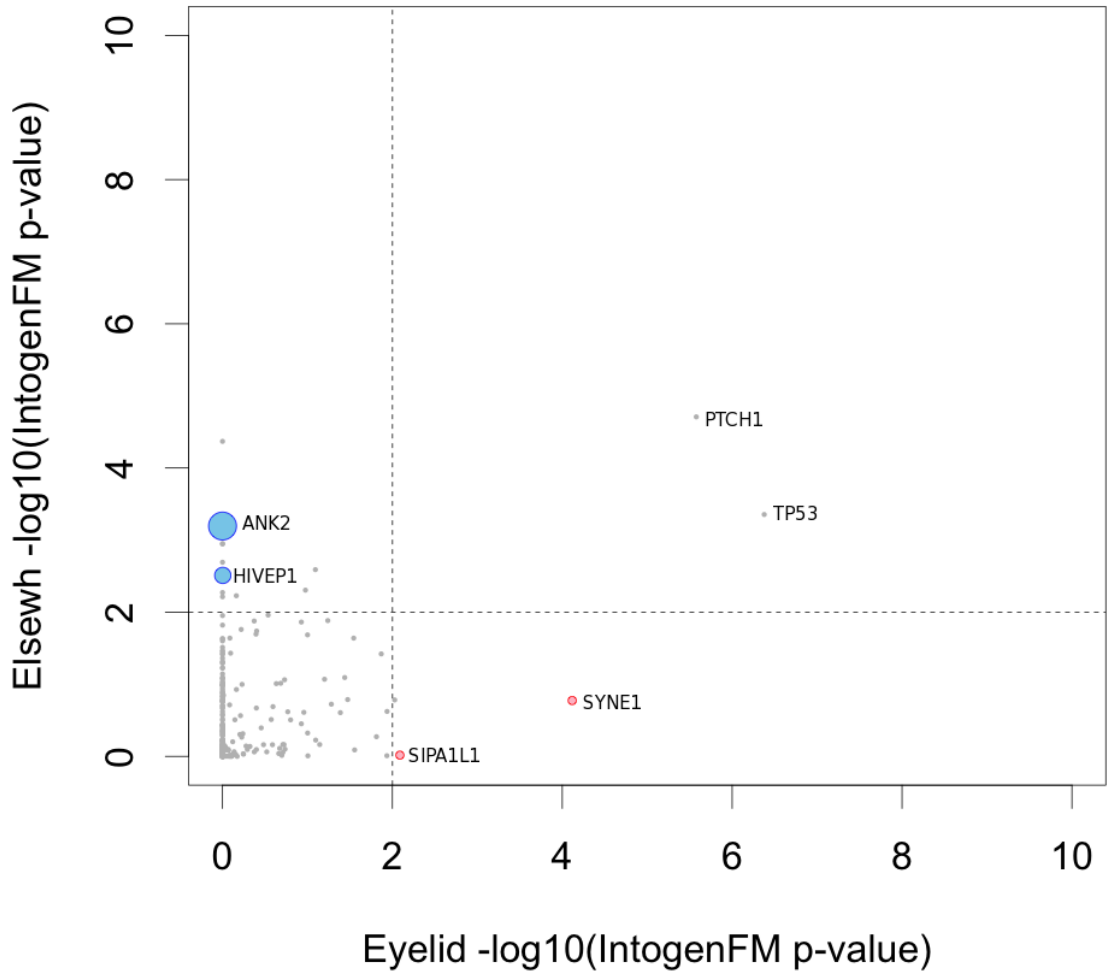


Figure 3.16 Randomised comparison of Intogen drivers to identify location specific genes in nodBCC. Drivers identified using Intogen were compared across the two localities, periocular and elsewhere, and a randomised comparison test performed to discover those genes that are more likely to picked by chance and those only found using WES discovery.

3.2.9 Stromal morphoeic BCC

Epithelial cancer behaviour is not restricted to the tumour borders delineated on microscopy; the surrounding stroma has been shown to play a key role in local spread.(Davidson et al., 2014) mBCC has long been associated with snake-like projections of cancer cells with surrounding inflammatory reaction containing fibroblast cells (Figure 3.17). This stromal soup could be beneficial to the tumour or simply be the body's immune response. Removal of the tumour alone is tricky and hence Mohs surgery is advocated to ensure histological removal, however, the leaving behind of the stromal reaction has not thought to be an issue. Local recurrence is a risk with mBCC and as this tumour rarely metastasis, one would think that this is due to the leaving behind of mutated cells. If the stroma contains mutations then this too may have to be removed along with tumour cells in future Mohs surgery.

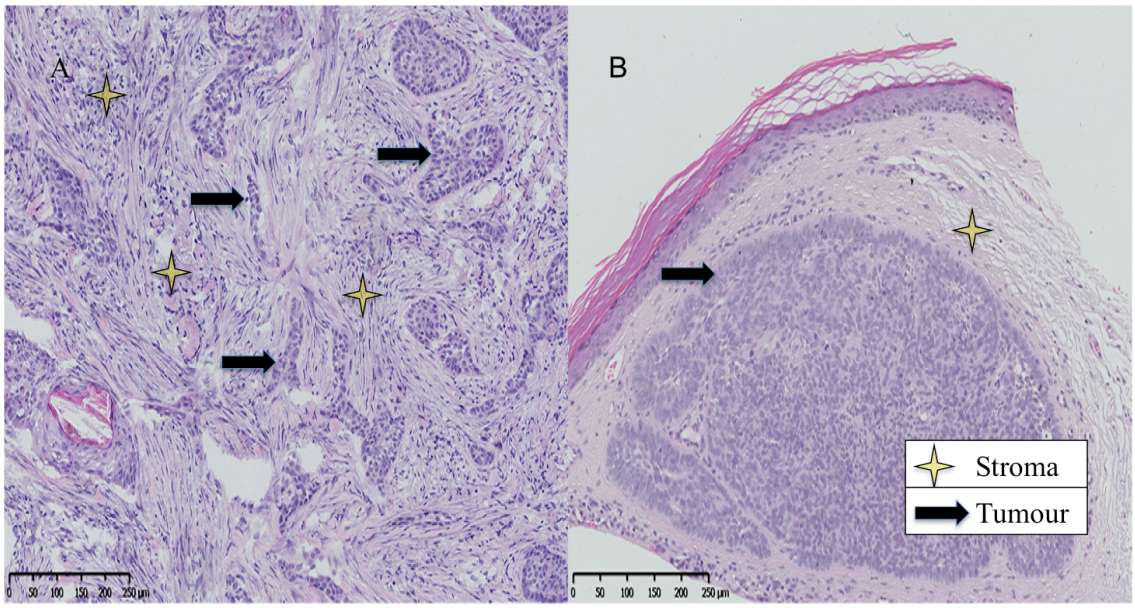


Figure 3.17 Histopathology of BCC subtypes. H&E staining of (A) Morphoeic BCC M4 highlighting the varied strands of tumour interdigitated in cellular stroma. (B) Nodular BCC for comparison demonstrating a well circumscribed tumour with relatively few cells in the surrounding stroma.

WES was carried out on both the tumour and histologically normal surrounding stroma of one microdissected mBCC sample. This revealed 150 shared nonsynonymous mutations (76.5% of all stromal mutations) including mutations in *ATR*, *EPHA3* and *Gli3*. Interestingly, the majority of the potentially pathological mutations revealed

identical variants (Table 3.20). The overall variant allele frequency (VAF), however, was low at 0.12 for the tumour at peak density and 0.04 for the stroma equating to a tumour purity of 24% and 8% respectively if assuming heterozygous change throughout the tumour (Figure 3.18). The highest VAF gene was *ROS1*, a proto-oncogene tyrosine kinase receptor and is discussed below.

	TUMOUR			STROMA		
Gene	**F**	**q/p**	**Driver**	**Stroma mBCC**	**SIFT**	**Poly**
TTN	8		I	c.57578 C>T c.22301 p.G19118E C>T p.G5759E		
EPHA3	3		I,CD	c.862 C>T p.P213S	0.00, D,L	0.96, P
ATR	2	0.487	I, HCD	c.4218 G>A pP1336S	0.17, T,H	0.85, P
C6	7	0.040	Mut	c.772 C>T pE170K	0.0, D,H	0.99, P
Gli3	1			c.2688 G>A p.S868L	0.01, D,H	0.999, P
ROS1	5		Dom	c.5628 G>A p.S1810F.	0.00, D, H	0.988, P
EPHA3	3		I,CD	c.862 C>T p.P213S	0.00, D,L	0.96, P
EPHB1	2			c.463 G>T p.E31D	0.36, T,H	0.98, P

Table 3.20 Driver mutations in mBCC tumour and stroma. Driver mutations in morphoeic tumour assessed for presence in the stroma underwent Intogen, SIFT and Polypen analysis. The top 12 genes were arranged using tumour Intogen using a fm-bias value <0.05. The mutations detailed above were the same one seen in the microdissected tumour. F=frequency, q=q-value, p=p-value, I=Intogen, Mut=MutSigCV, P=Probably, D=Damaging, T=Tolerated, L=low, H=High, SIFT, Poly=Polypen.

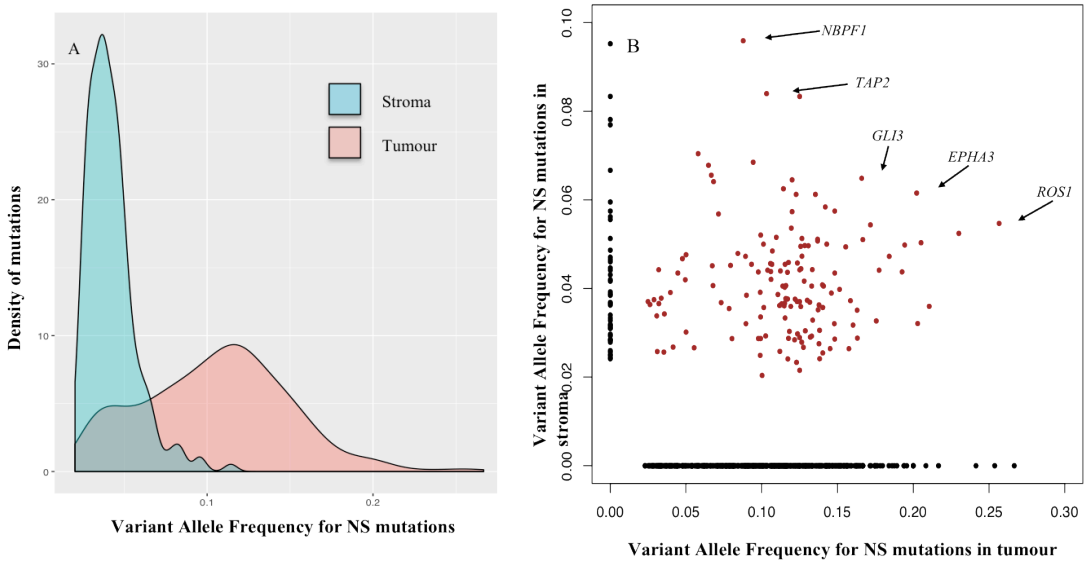


Figure 3.18 VAF in mBCC tumour and stroma. Variant Allele frequency (VAF) for non-synonymous (NS) mutations in morphoeic BCC M4 highlighting (A) peak density of mutational burden and (B) 150 shared mutations between stroma and tumour with the corresponding VAF.

3.2.10 Stromal morphoeic BCC molecular pathway analysis from WES data

Pathways that were altered in the tumour were also analysed to see if they were altered in the stroma too. Only 3 pathways contained variants and are listed in Table 3.23, however, these growth pathways do act in concert during embryogenesis so they may also be important in promoting the stroma milieu to encourage tumorigenesis.

Pathway	Gene	Variant	SIFT	Polpen
Hedgehog	Gli3	c.2688G>A / p.Ser868Leu	0.01, D, H	0.999, PD
Axonal G	EPHA3	c.862C>T / p.Pro213Ser	0.36 T, H	0.975, PD
	EPHB1	c.463G>T / p.Glu31Asp	0, D, L	0.996, PD
	SLITRK3	c.1762C>T / p.Glu494Lys		
Wnt	TBL1X	c.1446C>T / p.Ser269Phe		
		c.1447C>T / p.Ser269Ser		
	PPP3CA	c.2159 G>A / p.Ser492Phe	0.04, D, L	0.001, B

Table 3.21 Altered driver pathways in mBCC stroma. G=guidance; F=frequency, q=q-value, p=p-value, I=Intogen, Mut=MutSigCV, P=Probably, D=Damaging, T=Tolerated, L=low, H=High, SIFT= , Poly=Polypen;

DISCUSSION

3.3 Whole exome sequencing in basal cell carcinoma (BCC)

3.3.1 PTCH1 is shared mBCC driver using MutSigCV and Intogen

Patched 1 (*PTCH1*)– chromosome 9q22.32

As mentioned in the introduction, this is an inhibitory receptor of the Hh signalling pathway that represses smoothened (SMO) when no ligand is bound. On binding of the ligand, PTCH1 repression is released and SMO goes on to activate downstream targets such as Gli transcription factors. Inactivating mutations in *PTCH1* result in a variety of pathologies and its discovery in basal cell naevus syndrome paved the understanding of BCC tumorigenesis. Loss of *PTCH1* (by deletion or methylation) also occurs in oesophageal basaloid squamous cell carcinoma, keratocystic odontogenic tumours, cervical, colorectal, gastric and ovarian tumours to name a few.(Saito et al., 2015, Qu et al., 2015, Chung and Bunz, 2013, Zuo and Song, 2013, Musani et al., 2013, Chakraborty et al., 2015) Eight out of 10 morphoeic tumours contained *PTCH1* mutations and the majority were frame shift as summarised by Table 3.3. WES of BCC by Jayaraman et al revealed 9 out 12 tumours had a *PTCH1* mutation and they were mainly nonsense mutations, however, they failed to state the subtype of BCC analysed.(Jayaraman et al., 2014)

3.3.2 Drivers in mBCC identified using MutSigCV algorithm

Coiled-Coil Domain Containing 108 (*CCDC108*)– chromosome 2q35

This is a poorly defined protein that is highly expressed in the testis.(Imsland et al., 2012) SCOC (Short Coiled-Coil Protein), however, is involved in axonal growth in C.elegans and axonal growth may be a driver pathway in BCC.(Su et al., 2006) Coiled-coil interaction is an important component of DNA methylation-driven gene silencing and disruption of this coiling interaction can prevent methyl-CpG binding domain protein 2 (MBD2)-containing nucleosome remodelling and deacetylation (NuRD) complex (MBD2/NuRD) ability to silence genes.(Gnanapragasam et al., 2011) DNA methylation plays a role in cancer through transcriptional silencing of gene regulators. Nevertheless, it is not clear if CCDC108 has any relation to these coiling processes.

F-Box Protein 5 (*FBXO15*)– chromosome 18q22.3

FBXO15 contains an ubiquitin E3 ligase complex that is involved in the ubiquitination of p-glycoprotein/ABCB1 present on cancer cell surfaces that confers drug resistance.(Katayama et al., 2013) F-Box proteins also play an oncogenic role in colorectal cancer and loss of FBXO15 enhances vincristine resistance in fibrosarcoma cells.(Gong and Huo, 2015, Okita et al., 2007)

POM121 Transmembrane Nucleoporin-Like 12 (*POM121L12*)– chromosome 7p12.1

Little is known about this 1283 base pair gene. It is expressed in numerous cancers with a predilection for squamous cell carcinoma of the head and neck.(Bamford et al., 2004)

Olfactory Receptor, Family 5, Subfamily H, Member 2 (*OR5H2*)– chromosome 3q12.1

The largest family in the genome are the olfactory receptors whose role is in olfactory signalling and odorant binding.(Persuy et al., 2015) Although they are frequently detected in cancer their biological role may not be relevant.(Xu et al., 2015)

Keratin Associated Protein 12-3 (*KRTAP12-3*)– chromosome 21q22.3

Keratin associated proteins are large genes that play an essential role in the maintenance of the ridged hair shaft. An association with cancer has not been made, although in this case the origin of BCC may be the hair follicle where these proteins are expressed.

Complement component 6 (*C6*) – chromosome 5p13.1

The C6 protein is part of the complement cascade and is an integral part of the membrane attack complex (MAC) that causes cell lysis. Activation of the complement system is via three pathways, the classical, alternative and lectin pathways, however, they all converge at C3 that leads onto the final common pathway with the formation of the MAC via C5b-C9. Deficiency in C6 induces a susceptibility to Neisseria meningitides infection.(Moya-Quiles et al., 2013) Complement activation occurs in breast, gastric, lung and brain cancer.(Mamidi et al., 2015) Loss of this final pathway will interfere with the body's cancer defence, and induce complement resistance to monoclonal antibody immunotherapy mediated MAC attack.(Di Gaetano et al., 2003) Cancer clearing cells such as NK or T-cells are less efficient when complement activation products are missing.(Imai et al., 2007)

SYNE1 – Chromosome 6q25.2

SYNE1 is a large gene that codes for several isoforms including the large (>1000KDa) protein nesprin-1, a nuclear membrane protein involved in structural cellular organisation in muscle and the central nervous system, especially Purkinje cells.(Gros-Louis et al., 2007) Its role within the nuclear membrane may have a determining function for embryonic stem cell differentiation with an increase in facilitation for cells to differentiate.(Smith et al., 2011) Mutations in *SYNE1* results in neurodegenerative disease, cerebellar ataxia and motor neuron disease.(Izumi et al., 2013, Noreau et al., 2013) *SYNE1* is methylated in lung cancer and mutated in colorectal cancer.(Papadia et al., 2014, Tessema et al., 2008, Sjoblom et al., 2006) A non-synonymous SNV (rs2295190) in SYNE1 is associated with an increased risk of invasive ovarian cancer.(Doherty et al., 2010)

3.3.3 Drivers in mBCC identified using Intogen algorithm

Patched 1 (*PTCH1*)– chromosome 9q22.32

See details written above in section 3.3.1 and Table 3.3 for specific mutations.

Filamin B (*FLNB*)– chromosome 3p14.3

The filamin family trio are non-muscle actin-binding cytoplasmic proteins involved in scaffolding, but also interact with a variety of signalling proteins highlighting their importance beyond 3D structure.(Nakamura et al., 2011) They carry out cell signalling during embryogenesis and *FLNB* is involved in bone morphogenesis whereby mutations result in bone defects seen in spondylocarpotarsal syndromes.(Feng and Walsh, 2004) FLNB interacts with histone deacetylase (HDAC) to regulate vascular endothelial growth factor (VEGF) induced expression of HDAC7 target genes such as matrix metallopeptidase 10 (MMP10) and nuclear receptor subfamily 4, group A, Member 1 (NR4A1). Loss of FLNB inhibits vascular permeability, and in ovarian cancer cells, promotes MMP9-mediated tumour invasion and VEGFA-mediated vascular growth.(Su et al., 2013) In contrast, *FLNB* overexpression is related to spindle type extracellular cancer invasion seen in 3D collagen models.(Iguchi et al., 2015) Five out of ten samples contained a missense variant in the gene. The RNAseq data shows FLNB as being non-significantly (p=0.48) downregulated (-0.2 log2FC, expression

7.3), which may suggest a loss of function mutation being more likely in the morphoeic subtype.

Titin (*TTN*)– chromosome 2q31.2

TTN codes for cardiac and skeletal muscle protein and is the largest known protein. Mutations in this gene are seen in cardiomyopathy and autoimmune diseases, however, it is important to point out that virtually everybody has a variant in their *TTN* gene.(Chavanas et al., 2008) Despite this, it is still reported as associated with many cancers including the most common, breast and colorectal, and as mentioned, is one of the most mutated genes found in cancer.(Lips et al., 2015, Kim et al., 2015, Kim et al., 2013c)

Ubiquitin Specific Peptidase 4 (Proto-Oncogene) *USP4*– chromosome3p21.31

USP4 has been shown to counteract aberrant Wnt activation and therefore acts as a tumour suppressor gene. Its location on chromosome 3 is often lost in epithelial tumours where Wnt is dysregulated such as lung, breast, kidney and ovarian cancer.(Zhao et al., 2009) Overexpression, although less commonly reported, has been seen in adrenocortical cancer and can be involved in the aggressive nature of hepatocellular carcinoma cancer, which may mean USP4 has a cell specific role.(Velazquez-Fernandez et al., 2005, Heo et al., 2014) USP4 has recently been shown to suppress p53 and tumour necrosis factor-α (TNFα)-induced NF-κB activation preventing transcriptional, pro-apoptotic functions and inflammation respectively by deubiquitating HDAC2.(Li et al., 2015d) A novel role for USP4 is in its promotion of DNA repair by homologous recombination and deranged repair often occurs in cancer resulting in large chromosomal rearrangements.(Wijnhoven et al., 2015, Liu et al., 2015b)

Chromodomain Helicase DNA Binding Protein 3 (*CHD3*)– chromosome17p13.1

CHD3 is part of a protein complex called CHD3/nucleosome remodelling and deacetylase (NuRD) that carries out two distinct enzymatic chromatin organising functions, namely a remodelling ATPase and histone catalytic deacetylase subunits 1,2 (HDAC1/HDAC2). In addition, there are non-enzymatic subunits as part of the

complex that includes metastasis associated 1-3 (MTA1-3), retinoblastoma-binding protein 4 and 7 (RBBP4,7), and CpG binding domains 2 and 3.(Lai and Wade, 2011) It was initially described in the autoimmune connective tissue disorder dermatomyositis, which is a paraneoplastic condition, and this association with cancer prompted investigation into its role in oncogenesis. HDAC1/HDAC2 interacts with RB1, repressor element-1 silencing transcription factor Corepressor 1 (RCOR1) and SIN3 Transcription Regulator Family Member A (SIN3A) suggesting it has a key suppressor role in the cell cycle.(Denslow and Wade, 2007) More recently, it has been shown to maintain the silencing of tumour suppressor genes by occupying their promoters and functionally maintaining their hypermethylated, inactive state.(Cai et al., 2014) For example, the CHD3/NuRD complex recruits DNA methyltransferases and Chromobox Homolog 5 (CBX5) in the epigenetic repression of Cyclin-Dependent Kinase Inhibitor 1A (CDKN1A) by DNA methylation.(Choi et al., 2013) It is an essential cofactor in TWIST1 mediated repression of E-cadherin promoter, a hallmark of EMT, and TWIST1 mediated migration, invasion and metastasis.(Fu et al., 2011) HDAC1 inhibitors have been developed but possess significant side effects with minimal success.(Laugesen and Helin, 2014)

Erythropoietin-producing hepatocellular carcinoma (EPH) Receptor A3 (*EPHA3*)–chromosome 3p11.1

EPH receptors are members of the tyrosine kinase family and interact with two classes of ligands, A and B. Unlike other tyrosine kinase receptors that undergo dimerisation, EPH receptors require a minimum of four (tetramer) to form clusters, thus only a heterozygous loss would act in a dominantly negative fashion.(Janes et al., 2014) It preferentially binds the ligand ephrin-A5 and this complex interacts with metalloproteinase ADAM10. During embryogenesis, it plays a key role in patterning of the vascular and nervous system including axonal guidance. *EPHA3* is the most frequently mutated gene in lung adenocarcinoma and is a loss of function change.(Ding et al., 2008) Mutations are not only restricted to tumour tissue, but are also found in the surrounding stroma indicating a possible role in the tumour microenvironment.(Vail et al., 2014). *EPHA3* was shown to be a tumour suppressor gene in head and neck SCC and mutations have been identified in a variety of cancers including pancreatic, colorectal, and breast cancer.(Lee et al., 2010) In contrast, EPHA3 overexpression promotes aggressive behaviour and self-renewal in glioblastomas.(Day et al., 2013)

Depending on cellular context along with their bidirectional potential, it is important to determine their tissue specific role before embarking on therapeutic modification. Activating and inhibiting monoclonal antibodies have been designed against the EPH/ephrin system and antibodies can be radiolabelled to help imaging or targeted radiotherapy.(Barquilla and Pasquale, 2015) A recombinant, de-fucosylated human immunoglobulin G1κ (IgG1κ), KB004 (KaloBios Pharmaceuticals, California, USA), causes sustained EPHA3 activation resulting in tumour cell contraction and apoptosis and is currently being tested in phase 1 plus 2 clinical trials for haematological malignancies.(Boyd et al., 2014) Three out of ten samples had a variant, two missense and one nonsense.

NFS1 Cysteine Desulfurase (*NFS1*)– chromosome 20q11.22

Iron-sulphur clusters are part of proteins that are involved in a variety of cellular functions including respiration and gene expression.(Fosset et al., 2006) For the creation of iron-sulphur cluster, a complex involving NFS1-LYR Motif Containing 4 LYR (LYRM4) - Iron-Sulfur Cluster Assembly Enzyme (ISCU) is required.(Parent et al., 2015) The role it plays in cancer is not clear and is something that needs to be discerned.

Elongation Factor Tu GTP Binding Domain Containing 2 (*EFTUD2*)– chromosome 17q21.31

EFTUD2 is a GTPase component of a spliceosome complex that removes introns from pre-messenger RNA to form mature mRNA.(Staley and Woolford, 2009) It plays a fundamental role in embryogenesis and mutations in this gene cause mandibulofacial dysostosis with microcephaly.(Luquetti et al., 2013) EFTUD2 is also involved in the regulation of the innate immune system through alternative splicing of Myeloid Differentiation Primary Response 88 (MYD88) mRNA, a toll-like receptor signalling adaptor.(De Arras et al., 2014) In addition, it modifies the growth regulators (possible tumour suppressor genes) retinoic acid inducible gene 1 (RIG-1/DDX58) and melanoma differentiated protein 5 (MDA5/IFIH1) expression, which are positively correlated with EFTUD2.(Zhu et al., 2015, Tsai et al., 2009)

HECT Domain Containing E3 Ubiquitin Protein Ligase 4 (*HECTD4*)– chromosome 12q24.13

HECTD4 is a ubiquitin ligase family member that are involved in a plethora of cellular functions and they have been implicated in tumour growth and metastasis. Overexpression of E3 ubiquitin ligase promotes cell proliferation and survival via erythroblastic leukaemia viral oncogene homolog 2 (ErbB2) and epithelial growth factor receptor (EGFR) in breast and prostate cancer.(Chen et al., 2008) An intronic SNV in this gene (rs11066280) has been associated with a poorer outcome in oesophageal squamous cell carcinoma in a genome wide association study (GWAS), possibly due to alternative splicing creating a more potent transcript.(Zhang et al., 2015a, Pagani and Baralle, 2004) Four out of ten samples contained a variant, 2 missense and 2 nonsense.

Switch (SWI)/Sucrose Non-Fermentable (SNF) Related, Matrix Associated, Actin Dependent Regulator of Chromatin, Subfamily A, Member 4 (*SMARCA4*) – chromosome 19p13.2

Chromatin modelling can be regulated by complexes that either modify the histone tail or nucleosomes in an ATP fashion. These complexes play an essential role in gene expression and SMARCA4 is a member of the ATP dependent chromatin remodeler.(Wilson and Roberts, 2011) In addition, it plays a role in actin cytoskeleton with loss of function contributing to loss of actin cytoskeleton, loss of CD44, loss of ROCK1 expression and subsequent increase in invasion and motility.(Asp et al., 2002) Heterozygous knockout in mice (*smarca4 +/-)* are prone to epithelial tumours and highlights its role as a tumour suppressor gene.(Bultman et al., 2000, Rodriguez-Nieto and Sanchez-Cespedes, 2009) Frequent inactivating mutations are seen in lung cancer, rhabdoid tumours, pancreas, prostate and breast cancer.(Wong et al., 2000, Medina et al., 2005, Dal Molin et al., 2012). Furthermore, restoration of SMARCA4 reversed malignant phenotype in cell lines and inhibited growth. Loss of SMARCA4 causes a compensatory increase in expression of SMARCA2 as part of a complex, which has oncogenic properties.(Wilson et al., 2014) Mutations have been found in medulloblastoma, a tumour with similar associated aberrant signalling pathways and underscored by their occurrence in Gorlin syndrome, although this was only in 4 cases out of 92.(Dahlin et al., 2015, Pugh et al., 2012) *SMARCA4* was present in 4 out of ten samples so possibly has a higher frequency in mBCC. Immunostaining for SMARCA4 in BCC samples demonstrated a reduced protein expression along with SCC, but not a reduced RNA expression in SCC; BCC RNA levels were not reported.(Bock et al.,

2011) Furthermore, the BCC subtype was not mentioned. Four missense variants were found in four samples.

Folliculin Interacting Protein 1 (*FNIP1*)– chromosome 5q31.1

FNIP1 interacts with folliculin (FLN) and AMP kinase, both regulators of metabolism. Germline loss of function mutations in FLN results in an inherited kidney cancer syndrome (Birt-Hogg-Dube syndrome) defined by cutaneous fibrofolliculomas, pulmonary cysts, and kidney cancer. FNIP1 deficiency in mice (*fnip1-/-*) blocks B cell development resulting in immunodeficiency.(Park et al., 2012, Hasumi et al., 2015) Folliculin acts as a tumour suppressor gene via FNIP1-AMPK binding.(Baba et al., 2006) Activation of AMPK increases glycolysis and oxidative phosphorylation along with decreasing rapamycin mTOR driven cell growth.

Choline Kinase Beta (*CHKB*)– chromosome22q13.33

Two choline kinase enzymes (CHK A and B) catalyse the phosphorylation of choline in the biosynthesis of phospholipid phosphatidylcholine.(Wu and Vance, 2010) *CHKB* is a large gene of 22 Kb and plays an important role in bone embryogenesis and muscle function. Loss of function results in bone deformation and muscular dystrophy.(Mitsuhashi and Nishino, 2013, Sher et al., 2006, Li et al., 2014c) Transcription of *CHKB* is regulated by the protein kinase C pathway which represses the *CHKB* promoter.(Kuan et al., 2014) Dysregulated choline metabolism occurs in many cancers, but this has been mainly attributed to overexpression of CHKA with little known about the role of CHKB.(Glunde et al., 2015) More recently, it has been shown that the relative balance of CHKA to CHKB is more important for cancer cell survival: if CHKA is knocked out alone, mitosis can be halted and apoptosis occurs, but if both isoforms are knocked out, then cells no longer undergo cell death.(Gruber et al., 2012) Thus, the loss of *CHKB* confers an anti-apoptotic survival of cancer cells.

Rho GTPase Activating Protein 35 (*ARHGAP35*) – chromosome 19q13.32

The main action of ARHGAP35 is to convert Rho-GTP to Rho-GDP thereby inactivating the Rho pathway. By binding to p120, it forms a complex which positively regulates the Ras pathway by converting Ras-GTP to Ras GDP.(Arthur and Burridge, 2001) It also acts as a substrate for breast tumour kinase (brk) and phosphorylated ARHGAP35 plays an essential role in mediating the migratory and proliferative effects

of brk through Ras (activation) and Rho (inactivation).(Shen et al., 2008) Phosphorylated ARHGAP35 is regulated by the EGFR pathway in lung adenocarcinoma and involves invasion (via Ras) where is presence is associated with an aggressive phenotype.(Notsuda et al., 2013) Deliberate overexpression in the mouse model of pancreatic cancer resulted in a reduction of tumour invasion and metastasis.(Kusama et al., 2006)

CREB (c-AMP response element binding protein) binding protein (*CREBBP*) – chromosome 16p13.3

CREBBP binds specifically to phosphorylated CREB and acts as a transcriptional co activator. In addition, it has a histone acetyltransferase (HAT) domain, however, it does acetylate both histone and non-histone proteins. It is vital during embryogenesis and autosomal dominant loss of function mutations within this gene result in broad thumb-hallux syndrome (Rubinstein-Taybi syndrome). Increased frequency of tumours has been seen in this syndrome, in particular leukaemia and brain tumours. (Hutchinson and Sullivan, 2015) *CREBBP* mutations are seen in relapsed acute lymphoblastic leukaemia and in particular found within the HAT domain.(Mullighan et al., 2011)

3.3.4 Shared nodBCC drivers using MutSigCV and Intogen

Patched 1 (*PTCH1*)– chromosome 9q22.32

See section 3.3.1 for description of the gene in BCC. Six out of the ten-nodBCC samples had potentially pathological mutations in PTCH1. Two were frameshifts, two were nonsense and two were missense. In Gorlin syndrome, the majority of variants (65%) tend to be nonsense change, 16% missense, 13% splice site and 6% frameshift.

Tumour protein p53 (TP53) –chromosome 17p13.1

The archetypal tumour suppressor gene, TP53 has a major role in the stress response by carrying out DNA repair, cell arrest, senescence and apoptosis, all of which are defence mechanisms against tumour activity. *TP53* plays a cardinal role in transcription and is a master regulator of the cell cycle. It is not restricted to suppressing tumour initiation, but has activity in all stages of cancer progression, aggressiveness and metastasis highlighting the multistep process of malignant transformation. Loss of function is typically seen in many different cancer types by affecting the ability of the

TP53 protein to bind to DNA to activate transcription. A gain of aberrant advantageous tumour signalling function could also be happening and more difficult to decipher. Single nucleotide variant changes are most commonly seen and occur throughout the *TP53* coding sequence. Predilection occurs for exons 4 to 9, the area of the gene that codes for the protein DNA binding site domain. Furthermore, most mutations are heterozygous, meaning that both wildtype and mutant protein are present, however, they don't compete equally. Normal functioning TP53 requires the formation of a tetramer to carry out transcription which one mutant protein can disrupt, thereby acting in a dominantly negative fashion. Loss of heterozygous provides the final insult by inactivating any wildtype *TP53* allele. There is no denying that loss of wildtype or gain of oncogenic mutant function TP53 is a driver gene, but it is not clear what stage of tumorigenesis the effect is having. Inactivation of TP53 has been shown to occur in late adenoma early carcinoma transition for colorectal cancer; however, loss of the short arm of 17p has been shown in early adenoma. In pancreatic cancer, loss of TP53 occurs late. Ductal carcinoma in situ (early breast cancer) has demonstrated TP53 loss as has premalignant lesion Barrett oesophagus where the diploid cell is primed with loss of TP53. Environmental insults can induce TP53 mutations with smoking inducing a G to T transversion occurring at codons 154, 157, 158, 245, 248, and 273. UV induced skin cancers show early TP53 change each with their own individual hotspots. Normal eyelid skin accumulates *TP53* mutations suggesting a constant background mutational presence that may be normal for skin. Seven out of the ten samples had pathological *TP53* mutations, the same frequency as *PTCH1*.

Fatty acid desaturase 1 (*FADS1*) – chromosome 11q.22

The biosynthesis of unsaturated fatty acids requires the expression of dietary fat genes such as FADS1 which code for enzymes that add a double bond to a fatty acid chain.(Lee and Park, 2014) Genome wide association study has shown that the FADS1/2 locus is associated with long chain fatty acid, which in turn is risk to health.(Dorajoo et al., 2015) Furthermore, dysfunctional adipose tissue has been shown to be involved in carcinogenesis and cancer progression.(van Kruijsdijk et al., 2009) A large epidemiological study has identified FADS1 locus as a risk factor for colorectal cancer and decreased expression poses a poorer prognosis in oesophageal cancer.(Zhang et al., 2014, Li et al., 2015b) FADS1 mediates the formation of inflammatory mediators and is found to be downregulated in the tissue surrounding

prostate cancer, which is thought to aid local tumour invasion.(Ribeiro et al., 2012) FADS1 is downregulated (-0.6 log2FC, exp 5.3) in nodBCC but just non-significantly (P=0.09).

Protein Phosphatase, Mg2+/Mn2+ Dependent, 1D (*PPM1D*) – chromosome 17q23

This protein is a negative regulator of the cell stress response and is part of the PP2C family. It is an oncogene by the fact that is inhibits tumour suppressor pathways including TP53 via dephosphorylation.(Lu et al., 2005) PPM1D potentiates the Wnt/β-catenin pathway, which in turn increased the migration and invasion behaviour of pancreatic cancer cells.(Wu et al., 2016a) Increasing expression levels are seen in higher grades of meningioma suggesting it is more prevalent in aggressive tumours.(Fukami et al., 2016) Overexpression has been reported in hepatocellular, breast, ovarian, neuroblastoma and pancreatic cancer.(Ogasawara et al., 2015) Novel inhibitors have been developed against PPM1D including peptide and small chemical inhibitors. SL-176 inhibits PPM1D, which in turn suppresses proliferation of a breast cancer cell line via activation of TP53.(Ogasawara et al., 2015) Nevertheless, SL-176 also reduces growth in p53 null cells suggesting it suppresses division in a TP53 independent manner. As 7 out of 10 samples have a presumed loss of function TP53, if PPM1D is a gain of function change, then SL-176 could a frontrunner as a p53 independent inhibitor of cell growth in PPM1D overexpressed cancer states. Two out of ten had a PPM1D variant. However, RNA sequencing data shows PPM1D is non-significantly (P=0.9) downregulated (-0.1 log2FC, expression 3.2).

TAF4b RNA Polymerase II, TATA Box Binding Protein (TBP)-Associated Factor (*TAF4B*) – chromosome 18q11

TATA-box binding protein associated factors (TAFs) are part of fundamental transcription machinery and there are around 14 variants. TAF4 inactivation in the epidermis leads to hyperplasia in the mouse model and inhibits embryonic stem cell proliferation.(Fadloun et al., 2007) In contrast, TAF4B supports embryonic stem cell proliferation in cooperation with transcription factor POU Class 5 Homeobox 1 (POU5F1).(Bahat et al., 2013) TAF4B requirement for stem cell development and self-renewal has been elucidated in spermatogenesis (Oatley and Brinster, 2008) It is, however, part of the fundamental transcription machinery and shown to be aberrant in

several cancers. It is abnormal in 26% of high-grade ovarian cancer mainly as a gain of function or amplification.(Ribeiro et al., 2014) Knockdown, however, promotes migration of colon cancer.(Kalogeropoulou et al., 2010) Expression of TAF4B is induced by MYC in a non-canonical fashion along with their co-expression in leukemic cells.(Teye et al., 2008) TAF4B is non-significantly (P=0.7) down regulated (-0.2 log2FC, exp 3.0) in nodBCC.

Thyroid Hormone Receptor Interactor 4 (TRIP4) – chromosome 15q22

This is a transcription coactivator that interacts with many transcription factors including cAMP-response element binding protein, binding protein (CREBBP), E1A Binding Protein P300 (EP300) and Nuclear Receptor Coactivator 1(NCOA1). The driver gene EP300 is altered in 2 out of 10 nodular and NCOA1 is altered in one tumour. In concert with its binding protein, Activating Signal Cointegrator 1 Complex Subunit 1 (ASCC1), it may have a role in putative cross talk with JUN and NF kappa B linking inflammatory which may be important in oesophageal cancer.(Orloff et al., 2011) Ufmylation (by Ubiquitin-Fold Modifier, UFM1) of TRIP4 is an essential step in the development of oestrogen receptor positive breast cancer. TRIP4 is non-significantly (P=0.9) up regulated (0.1 log2FC, exp 5.1) in nodBCC.

3.3.5 Drivers identified using MutSigCV algorithm

Gamma-Aminobutyric Acid (GABA) A Receptor, Alpha 6 (*GABRA6*) - chromosome 5q34

This is a receptor that mediates inhibitory action of GABA and is mainly associated with brain. Altered genetic function is therefore associated with schizophrenia, epilepsy, panic disorders and other behavioural changes.(Hirose, 2014) However, it is also involved in AKT/PK3B and apoptosis pathways. It prevents neural apoptosis via the JNK3 pathway, but has been described to have actions outside the central nervous system by preventing apoptosis in pancreatic islet cells.(Prud'homme et al., 2014) It also has inhibitory action via NF-kB, an activator of the immune response.(Prud'homme et al., 2013). Both these actions would confer an advantage to the tumour. GABA A receptors have shown to stimulate basal-like breast cancer in cell lines, promote growth of hepatocellular cancer and the growth of pancreatic cancer.(Li et al., 2012c, Sizemore et al., 2014, Takehara et al., 2007)

Olfactory Receptor, Family 52, Subfamily J, Member 3 (OR52J3) – chromosome 11p15.4

OR52J3 is an odorant receptor involved in smell transmission. It is part of a large group of proteins that biologically (with the evidence known) should not have any relation to cancer. A polymorphism in this gene, rs17350764, has been associated with asthma in a genome-wide association study (GWAS), however, they do admit it is difficult to understand the functional significance.(Song and Lee, 2013)

Testis specific, 13 (*TSGA13*) – chromosome 7q32.2

TSGA13 is abundantly found in testis; however, it is also expressed in normal tissue including spleen, stomach, bladder, and thyroid. In addition, it is a gene that undergoes genomic imprinting, a process whereby one of the parental copies are silenced in the germline, so that there is only monoallelic expression in somatic cells. This phenomenon, which only occurs in around 100 genes in humans, makes it particularly venerable to genomic alterations such as those seen in cancer and more likely to have a functional effect. These genes tend to be conserved across species and are thought to have vital roles especially during embryogenesis. Reduced expression has been shown in liver, breast and stomach cancers and so it is possible that loss of function in this gene confers a benefit to the tumour.(Zhao et al., 2015a)

Tripartite Motif Containing 39 (*TRIM39*) – chromosome 6p22.1

It has shown to be an important factor for inflammatory signalling and negatively regulates the NFkappaB pathway, a pivotal player in cell survival, immunity, inflammation, carcinogenesis and cell growth.(Suzuki et al., 2016) Cell survival action occurs through the regulation of the cell cycle through p21 and this interaction has been shown in hepatocellular carcinoma.(Zhang et al., 2012) TRIM39 stimulates the innate immune system and provides protection against virus.(Wang et al., 2016b)

Chromosome 10 Open Reading Frame 82 (*C10orf82*) – chromosome 10q25.3

Malignant melanomas, testicular cancer and a sample of BCC have shown high expression of C10orf82.(Uhlen et al., 2015)

Fatty acid desaturase 1 (*FADS1*) – chromosome 11q.22

See section 3.2.6i

3.3.6 Drivers identified using Intogen algorithm

Collagen, Type IV, Alpha 4 (*COL4A4*) – chromosome 2q36.3

Basement membranes are ubiquitous across the body, however, this gene codes for collagen IV in only a subset of basement membranes making it more tissue specific. Mutations in this gene are found in hereditary glomerulonephropathy (Alports syndrome) and familial benign haematuria. Patients with Alports syndrome and loss of the 5-prime end of the *COL4A4* gene can develop leiomyomatoma, a smooth muscle tumour that is essentially indolent, transforming to a malignant tumour in only 0.1%) (Zhang et al., 1996) Downregulation of COL4A4 was shown to occur in oesophageal cancer.

Rho GTPase Activating Protein 35 (*ARHGAP35*) – chromosome 19q13.3 see section 3.2.3, p92

This driver gene was also mutated in the mBCC subtype and so is a common driver (trunk mutation) that may be important in early tumorigenesis.

5-Methyltetrahydrofolate-Homocysteine Methyltransferase (*MTR*) – chromosome 1q43

The final step in the synthesis of methionine is catalysed by MTR. There has been some interest in folate levels and polymorphisms in MTR; however, no relationship was seen with MTR variations in breast cancer, non-Hodgkin's lymphoma, or bowel cancer.(He et al., 2015, He et al., 2014, Taflin et al., 2014)

v-myc Avian Myelocytomatosis viral oncogene Neuroblastoma Derived Homolog (MYCN) – chromosome 2p24.3

MYCN codes for a nuclear protein that requires dimerisation in order to activate DNA transcription. As the name suggests it is commonly found in neuroblastoma. MYCN is frequently mutated in medulloblastoma, another Hh pathway driven tumour, however, a gene risk study on 243 children with medulloblastoma did not show an increased risk for a particular *MYCN* SNV.

Asparaginyl-TRNA Synthetase (*NARS*) – chromosome 18q21.31

Although this is traditionally known as a tRNA enzyme responsible for esterification of a cognate amino acid to its tRNA as part of protein synthesis, it has also been

associated with growth promoting function. FGF2 induced NARS promotes the survival of osteoblasts through PI3K/Akt signalling.(Park et al., 2009).

Other driver genes picked from Intogen analysis

Protein Phosphatase, Mg2+/Mn2+ Dependent, 1D (*PPM1D*) – chromosome 17q23

See section 3.2.5

AT Rich Interactive Domain 1A (SWI-Like), (*ARID1A*) – chromosome 1p36

This gene is part of the SWItch/Sucrose Non-Fermentable (SWI/SNF) chromatin remodelling complex and plays a role in regulating transcription of gene, which are repressed by surrounding chromatin. As a consequence, it is thought to be a tumour suppressor gene its loss could potentially be a prognostic marker. Histone H2B is a main component of chromatin and thought to be a target of ARID1A.(Li et al., 2010, Wu et al., 2016b) Loss of ARID1A leads to hypermethylation of MLH1 mismatch repair gene among others. In addition, ARID1A negatively regulates telomerase reverse transcriptase (TERT), hence loss of function confers a survival advantage to cancer cells.(Suryo Rahmanto et al., 2016) Mutations are most commonly seen in ovarian clear cell carcinoma, endometrial cancer, renal clear cell carcinoma and gastric carcinoma.(Takeda et al., 2016, Sato et al., 2016) Gastric cancer cases with partial or reduced loss had a minimal disease free survival effect, but complete loss had a worse disease free survival conferring a hazard ratio of 1.7. (Kim et al., 2016) Low levels were seen in 71% of 432 cases of breast cancer and were an independent predictor of poor disease free and overall survival.(Cho et al., 2015) In colorectal cancer, loss of function was seen in 8% of cases and although it was associated with poor differentiation, lymphovascular invasion and higher tumour stage, it did not correlate with prognosis.(Lee et al., 2016) Similar findings occurring in hepatocellular cancer where low expression correlated with larger tumour size, but not prognosis.(Zhao et al., 2015b) Two out of ten samples have a variant in the gene. ARID1A is non-significantly (P=0.5) down regulated (-0.2 log2FC, expression 7.4) in nodBCC.

Interferon Regulatory Factor 2 (*IRF2*) – chromosome 4q35

As a member of the transcription regulatory family it can both inhibit and activate transcription factors. Originally described to regulate the beta interferon gene,

downregulation of IRF2 on human leukemic cells resulted in arrested growth highlighting its oncogene properties.(Choo et al., 2008)

3.3.7 Drivers compared to external BCC cohort

A genomic analysis of a large cohort of 293 BCC revealed 7 significantly mutated genes using MutSigCV on 121 of the samples with a q-value of <0.1. (Bonilla et al., 2016). However, the majority underwent targeted sequencing; 38 nodular WES and 8 morphoiec BCC only. These seven genes were assessed in our cohort of patients to identify any commonality (Table 3.22).

Bonilla et al	MBCC	NodBCC	Total our cohort (%)
PTCH1 (73%)	8	7	15 (75%)
TP53 (61%)	2	7	9 (45%)
SMO (20%)	1	2	3 (15%)
MYCN (30%)	1	4	5 (25%)
PTPN14 (23%)	1	1	2 (10%)
RPL22	1	3	4 (20%)
PPIAL4G	1	1	2 (10%)

Table 3.22 Presence of external BCC drivers within our cohort. Seven significant driver mutations with a q value of <0.1 using MutSigCV on a 293 cohort (sporadic, inherited, Vismodegib naïve and treated, all histological subtypes including unknown) that underwent WES and target sequencing compared to our cohort of 20 BCC.

The Bonilla cohort was further scrutinised to select for sporadic, morphoeic subtype BCC analysed using WES (as opposed to a cancer gene panel of a selected number of genes). After filtering the mutation summary table of 293 cases (supplementary table 1b of Bonilla *et al*) for WES, Vismodegib naïve and sporadic BCC revealed only 8 morphoeic and 38 nodular samples.(Bonilla et al., 2016) It may therefore be prudent to state that the aforementioned 7 BCC driver mutations reflect a nodBCC makeup as there sample size is heavily weighted to this histological subtype.

3.3.8 Drivers within morphoeic stroma

GLI Family Zinc Finger 3 (*GLI3*) – chromosome 7p13

Gli3 is part of the Hh signalling pathway and often called a transcriptional repressor. An autosomal dominant loss of function mutation in Gli3 leads to Grieg's cephalopolysyndactyly syndrome, a variable syndrome resulting in abnormal development of the limbs or head. Wnt activity controls Gli3 expression via β-catenin/Tcf pathway with the Gli3 gene revealing Tcf binding sites as part of its code.(Alvarez-Medina et al., 2008) Reducing the expression of Gli3 inhibits colonic cancer cell growth suggesting an antagonist of Gli3 would be beneficial to suppress cancer.(Trnski et al., 2015) Moreover, Gli3 expression is associated with cyclopamine (SMO inhibitor) resistance in pancreatic cancer cells and is tumorigenic as per colonic cancer cells.(Steg et al., 2010) Gli3 is non-significantly (P=0.6) upregulated (log2FC 0.4, exp 5.8) in morphoeic stromal tissue. The VAF in the stroma was low at 8%, however, it was low in the tumour at 26%.

Erythropoietin-producing hepatocellular carcinoma (EPH) Receptor A3 (*EPHA3*)– chromosome 3p11.1

See section 3.2.3. EPHA3 is non-significantly (P=0.5) upregulated (log2FC 0.8, exp 1.8) in morphoeic stromal tissue.

ATR Serine/Threonine Kinase (*ATR*) – chromosome 3q23

This belongs to the phosphatidylinositol 3-kinase (PI3K) family and *ATR* plays a vital role in the cell cycle by phosphylating checkpoint kinase 1 (CHK1), checkpoint proteins RAD17/9 and BRCA1 tumour suppressor gene. It is particularly activated during times of genotoxic UV light exposure where it causes arrest of the cell cycle, DNA repair and apoptosis.(Al-Khalaf et al., 2012) It shares a similar function with Ataxia Telangiectasia Mutated Serine/Threonine Kinase (*ATM*), however, is generally activated during single strand DNA damage whereas ATM is involved in double strand breaks. An autosomal dominant germline mutation in ATR can lead to Familial Cutaneous Telangiectasia and Oropharyngeal Predisposition Cancer Syndrome.(Tanaka et al., 2012) Interestingly, Imiquimod treatment of BCC activates TP53 dependent apoptosis by ATR pathway, thus it is important that the gene is

functioning for this medication to be effective.(Huang et al., 2016b) ATR is non-significantly (P=0.9) down regulated (log2FC -0.02, exp 5.7) in morphoeic stromal tissue.

Complement component 6 (*C6*) – chromosome 5p13.1

See section 3.2.4ii. Seven out of 10 mBCC tumour contain a variant.

ROS proto-oncogene 1 (*ROS1*) – chromosome 6q22

This is a tyrosine kinase receptor member of the insulin receptor superfamily. IT has been identified in cancer states, but its wild type function remains elusive. Chromosomal rearrangement with the formation of aberrant fusion genes have been reported with ROS1 in gastrointestinal, lung adenocarcinoma and inflammatory myofibroblastic tumours.(Davies and Doebele, 2013, Bergethon et al., 2012) The latter has similar characteristics to mBCC with an inflammatory fibroblastic response, local recurrence potential and minimal risk of metastasis. TFG-ROS1 and YWHAE-ROS1 tend to occur in ALK negative myofibroblastic tumours. ROS1 fusions seem to congregate to a particular subset of NSCLC, adenocarcinoma. Furthermore, there seems to be a loss of the non-rearranged ROS1 allele too. The ROS1 mutation with the mBCC tumour and stroma is also named in COSMIC (id 139667). Treatment of ALK positive NSLC with specific antibodies has been developed including Crizotinib (Xalkori®, Pfizer) and some have activity against ROS1. Resistance can occur to these drugs, particularly in the tyrosine kinase section of the oncogene. Interestingly, these drugs have affinity for ROS1 rearrangements too.

EPH Receptor B1 (*EPHB1*)- chromosome 3

This is an ephrin receptor and also part of the axonal guidance pathway. See section 3.2.4. Four out of 10 mBCC tumour contain a variant.

The mutational load of the morphoeic stroma is significantly high with 196 potentially pathogenic non-synonymous changes, ten of which are stop-gain changes. Eighty-one were probably damaging using the Polypen and SIFT algorithms. Shared potentially pathogenic variants using the two algorithms were 54 including *ROS1, ORSC1, PCDHBG1, ZNF381* and *VPS13C* that were also found in COSMIC. Potential drivers

within the stroma were determined using knowledge of their biological along with their cancer association (Table 3.22).

Gene	Protein Variant	Cancer association	Freq
PTPN22	p.Trp244*	Uterine leiomyomas, CLL	1
ARHGEF28	p.Tyr1232*		3
ASAP1	p.Gln318*	Invasive phenotype of laryngeal SCC, ovarian	2
SFMBT2	p.Gln596*	Regulation of prostate cancer cell growth	1
MYBPC1	p.Arg318*		2
XPO4	p.Gln412*	Tumour suppressor gene. Lost in liver cancer	1
CPNE7	p.Gln122*	Inflammation amplifier	1
CDH20	p.Trp684*	Early stages of tumorigenesis	3

Table 3.23 Potential drivers in stromal mBCC. Top 8 stop-gain non-synonymous changes. * stop/gain; F=frequency, q=q-value, p=p-value, I=Intogen, Mut=MutSigCV; Freq, frequency of gene variant in the morphoeic tumour cohort

CONCLUSION

3.4 Summary of genetic variants in BCC

BCC contains a large number of mutations many of which are not significant to tumour initiation, growth, progression or metastasis. Normal skin contains many mutations too and to identify the driver genes within the tumour required several approaches. WES underlined the huge tumour burden with the vast majority of NSSNV being UV induced (UV signature). Interestingly the location of the tumour was not related to tumour burden nor was the UV signature profile. Morphoeic is a more aggressive tumour, but it did not display any greater mutational burden than nodBCC nor did it have a greater UV signature. Driver genes within mBCC revealed a large number with little correlation between the two algorithms (Intogen and MutSigCV) except for PTCH1. This may represent the heterogeneous nature of this tumour and that it is fairly uncommon to find pure mBCC on histology without nodBCC or another subtype being present too. Pathway WES analysis revealed the importance of Hh signalling, which was supported by the RNAseq and protein expression (See chapter 4). Other significantly changed pathways includes up regulation of the axonal guidance pathway, down regulation of both the NK and FC epsilon pathways.

In contrast, nodBCC algorithm analysis revealed six common drivers and *TP53* is an integral change in tumour development. Pathway WES analysis places BCC pathway as number one and is the only upregulated pathway along with mismatch repair pathway on RNAseq analysis; Hh pathway was not feature on RNAseq analysis (see Chapter Four).

The skin stroma lies beneath the epithelium and consists of a superficial lymphoid and deep fibrous layer. WES of mBCC stroma shared 150 genes with its tumour counterpart including the identical mutation in *ATR*, *EPHA3* and *Gli3*, although the overall VAF was low in both samples. *EPHA3* is thought to act as a tumour suppressor gene and is the most frequently mutated gene in lung adenocarcinoma as a loss of function change.(Ding et al., 2008) Mutations are not only restricted to tumour tissue, but are also found in the surrounding stroma indicating a possible role in the tumour microenvironment.(Vail et al., 2014). The highest mBCC stromal VAF gene was a *ROS1*, a proto-oncogene tyrosine kinase receptor. This could be true stromal mutational burden or just undetectable cancer cells intermixed in the stromal soup that are being analysed. Either way it is clinically relevant when it comes to excision whereby the removal of this inflammatory material during Mohs micrographic surgery should be considered, as this may be the reason for local recurrence.

The vast majority of mutations in cancer are heterozygous, but this does not mean that they no effect if the wildtype allele is present.(Rentoft et al., 2016) However, knowledge of the copy number aberration would help to improve the interpretation of the potential driver gene and allow the focus on those which may have undergone LOH or amplification. In addition, functional studies to determine their true nature is required and allow the identification of drug targets; potential drug targets highlighted in this chapter are discussed further in Chapter Four.

CHAPTER FOUR

GENE EXPRESSION IN BASAL CELL CARCINOMA

4.1 Further aims of the BCC study

Genetic variants alone are not enough to discern the effects that are ensued by the tumour although a wealth of information that hints to possible biologically relevant changes has been documented in the last chapter. Looking at the gene expression in terms of RNA and protein is outlined in this chapter (Figure 4.1) to see if it complements the variance data, but also to identify those pathways that are modified but don't possess a mutation.

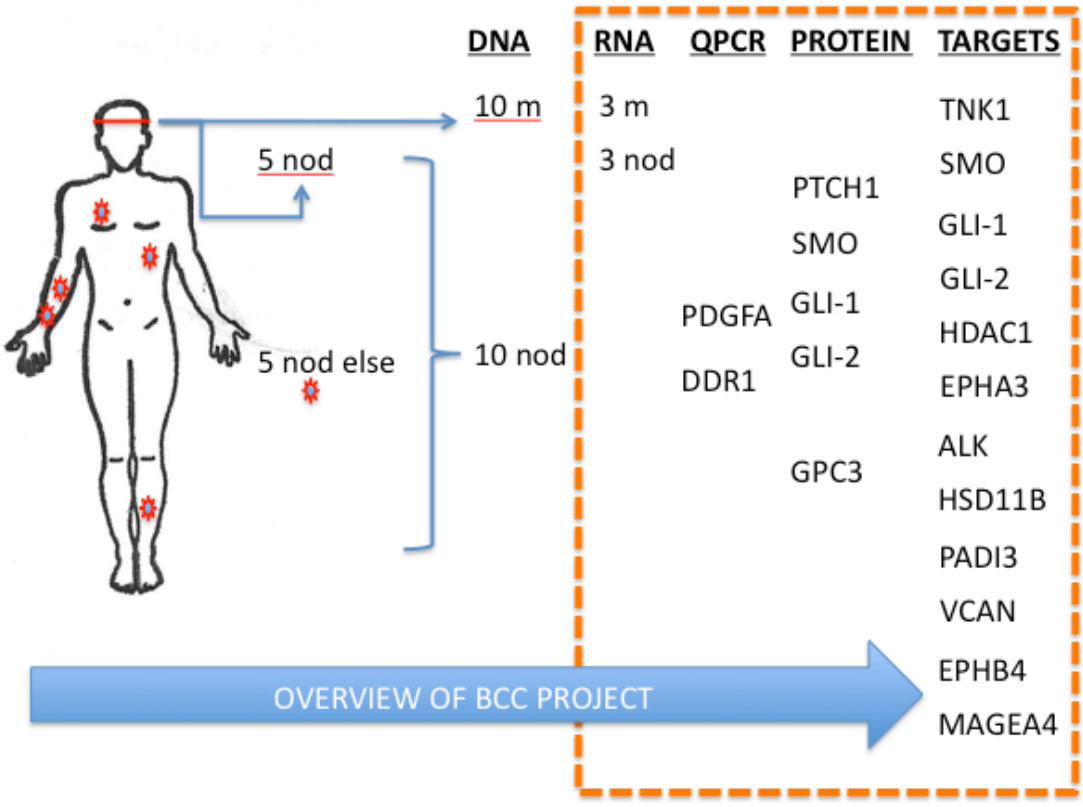


Figure 4.1 BCC project part 2. Overview of the BCC project highlighting the stepwise progression starting with genetic variant detection in the DNA, followed by genetic expression using RNAseq and validation using quantitative PCR plus protein expression. The orange box highlights the focus of chapter four. Nod, periocular nodular BCC; nod else, nodular BCC not within the high-risk H zone, m, periocular morphoeic BCC.

RESULTS

4.2 RNA sequencing in periocular BCC

4.2.1 Clinical and histological features of periocular BCC patients

Six BCC patients who underwent WES also had RNA sequencing of their tumour. This was compared against normal tissue at a large distance away from the cancer. Three were morphoeic and three were nodular BCC (Table 4.1). Each sample underwent fresh, frozen laser capture microdissection to ensure purity of sample. All 12 samples (6 tumours and 6 normal stoma) were run on a single lane (13 million reads each) and underwent the same processing conditions such as library preparation.

Tumour	Age	Gender	Location
M1	85	M	RLL
M2	47	F	LLL
M3	78	F	LLL
Nod3	71	M	RMC
Nod4	72	F	LLL
Nod5	46	F	RLL

Table 4.1 Clinical features of RNAseq samples. Samples used where also ones that underwent WES. M, morphoeic BCC; Nod, nodular BCC.

4.2.2 Periocular morphoeic BCC RNA expression

Raw reads counts were derived from processing high-throughput sequencing data through HTSeq (Python library) against Ensembl annotation.(Anders et al., 2015) A paired analysis of morphoeic tumour versus morphoeic stroma using limma R/Bioconductor software package (voom function) was preformed using raw p value less than 0.01 and log2FC greater than 1 or less than -1.(Ritchie et al., 2015) Downregulation in the tumour is defined as log2FC<-1 and upregulation in the tumour as log2FC > 1. Paired analysis of morphoeic tumour versus stroma revealed 288 differentially expressed genes using a P value < 0.01 and log2FC >1 or <-1. Log2FC ranged from -7.1 to 5.0 and expression levels between -1.8 to 9.1. A heatmap of the findings is summarised in Figure 4.2 and almost 60% (169 genes, blue colour) were downregulated and the vast majority, 143, being -2 log2FC or more.

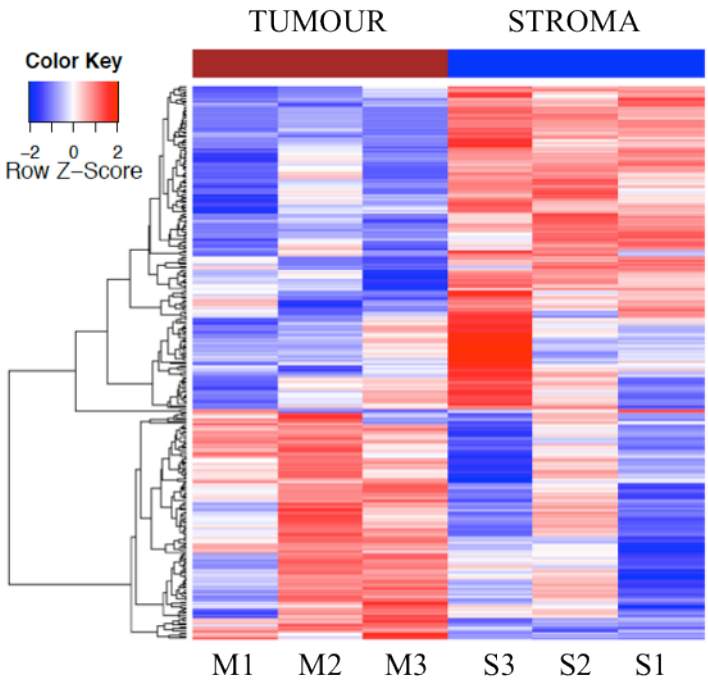


Fig 4.2 Heatmap of mBCC tumour and stroma. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression.

4.2.2.i Upregulated genes in morphoeic BCC

Genes found to be significantly upregulated were organised according to their Log2FC (Table 4.2) and then according to the highest expressed (Table 4.3). Selected genes of interest are further discussed in terms of their biological relevance and cancer association. Genes equally expressed in both tumour and stroma may not be picked up.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
AMER2	2.2	2.9	<0.001	HHIP	3.7	2.3	<0.001
ALK	3.8	2.8	<0.001	EDN2	3.3	2.3	<0.001
HSD11B1	3.2	2.8	0.002	CALCB	3.3	2.2	<0.001
PADI3	2.1	2.8	0.006	GPC3	4.2	2.1	0.001
TSHR	3.9	2.7	0.004	BOK	3.1	2.0	0.003
STAC2	6.0	2.4	<0.001	MAGEA4	2.1	1.9	<0.001

Table 4.2 Highest Log2FC upregulated genes in mBCC. Top 10 differentially expressed RNA are arranged according to the highest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
STAC2	6.0	2.4	<0.001	CALCB	3.3	2.2	<0.001
GPC3	4.2	2.1	0.001	EDN2	3.3	2.3	<0.001
TSHR	3.9	2.7	0.004	HSD11B1	3.2	2.8	0.003
ALK	3.8	2.8	<0.001	BOK	3.1	2.0	0.003
HHIP	3.6	2.3	<0.001	AMER2	2.2	2.9	<0.001

Table 4.3 Most upregulated genes in mBCC arranged according to expression levels. Top 10 differentially, highest expressed, upregulated RNA with a log2FC of greater than 2.0. Exp=expression, P-val=P-value.

4.2.2.ii Downregulated genes in morphoeic BCC

As mentioned, the majority of genes are downregulated in mBCC. Genes found to be significantly downregulated were organised according to their Log2FC (Table 4.4) and then according to the highest expressed (Table 4.5). Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
KRT13	3.1	-7.1	0.007	IGLV1-40	2.9	-4.2	<0.001
KRT4	3.0	-5.8	<0.001	IGKV3-11	2.9	-4.2	0.007
TGM2	2.5	-5.0	0.009	TMPRSS4	3.8	-4.1	0.003
CXCL17	2.5	-4.9	0.004	IGLC2	5.3	-4.0	P<0.001
PIGR	3.3	-4.7	0.009	IGHA1	5.7	-4.0	0.002

Table 4.4 Log2FC downregulated genes in mBCC. Top 10 differentially expressed RNA are arranged according to the lowest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
S100A8	8.7	-3.1	0.005	SERPINB1	6.8	-2.3	0.004
IGHG1	8.4	-2.7	<0.001	SLPI	6.5	-3.6	0.008
KRT6A	7.9	-2.5	0.001	IGJ	6.4	-3.3	<0.001
FABP5	7.9	-2.7	0.002	IGHA1	5.6	-4.0	0.002
IGLC3	7.0	-3.5	<0.001	IGHG4	5.3	-3.2	<0.001

Table 4.5 Most downregulated genes in mBCC arranged according to expression levels. Top 10 differentially, highest expressed, downregulated RNA with a log2FC of less than -2.0. Exp=expression, P-val=P-value.

4.2.3 Periocular nodular BCC RNA expression

Two hundred and eighty-eight genes were differentially expressed when nodular and stromal tissue were compared. Raw reads counts were derived from processing high-throughput sequencing data through HTSeq (Python library) against Ensembl annotation.(Anders et al., 2015) A paired analysis of nodular tumour versus nodular stroma using limma R/Bioconductor software package (voom function) was performed using raw p value less than 0.01 and log2FC greater than 1 or less than -1.(Ritchie et al., 2015) Downregulation in the tumour is defined as log2FC<-1 and upregulation in the tumour as log2FC > 1. A heatmap of the findings is summarised in Figure 4.3 and just under half (145) of the differentially expressed genes were down regulated and the range of log2FC was from -5.5 to 4.8 with expression levels between -2 to 11.

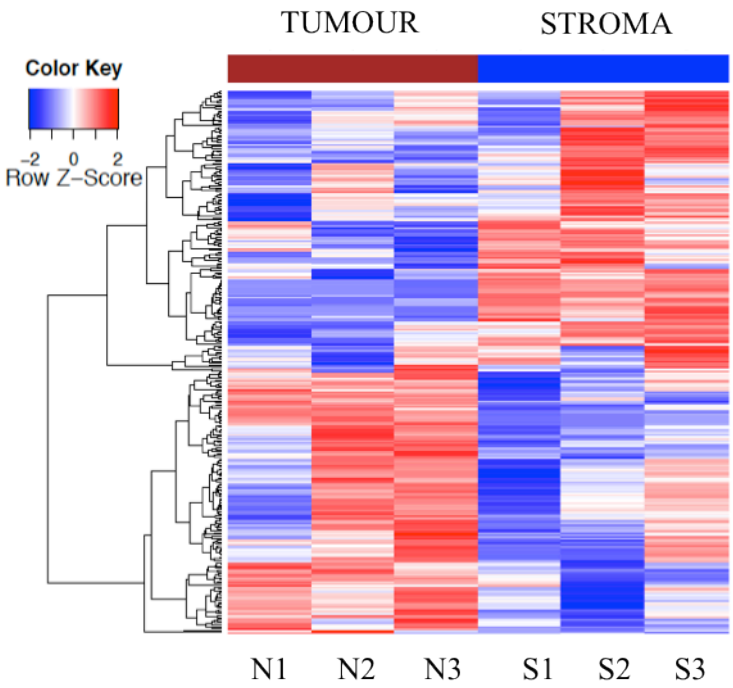


Fig 4.3 Heatmap of nodBCC tumour and normal stroma. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression.

4.2.3i Upregulated genes in nodular BCC

Genes found to be significantly upregulated were organised according to their Log2FC (Table 4.6) and then according to the highest expressed (Table 4.7). Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Exp	Log2FC	P val	Gene	Exp	Log2FC	P val
CAPN6	2.7	3.7	0.008	RP11-996F15.2	2.0	3.9	0.006
miR-3117	2.2	3.5	0.009	WIPF3	4.2	2.9	0.004
ZNF418	2.1	3.3	0.009	ZNF737	2.8	2.7	0.004
TLL1	2.0	3.2	0.003	TSHR	3.9	2.7	0.001
RP11-119F7.5	2.2	3.0	0.002	SHISA2	5.1	2.3	0.005

Table 4.6 Highest Log2FC upregulated genes in nodBCC. Top 10 differentially expressed RNA are arranged according to the highest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value

Gene	Exp	Log2FC	P val	Gene	Exp	Log2FC	P val
ABI3BP	8.3	2.2	<0.001	SLITRK2	4.9	2.2	0.001
STAC2	6.0	2.2	0.002	PLAG1	4.8	2.3	0.002
LMO3	5.3	2.0	<0.001	WIPF3	4.2	2.9	0.004
SOX11	5.3	2.3	<0.001	ZNF251	3.9	2.6	0.001
SHISA2	5.1	2.3	0.005	TSHR	3.9	2.7	0.001

Table 4.7 Highest expressed genes in nodBCC. Top 10 differentially, highest expressed, upregulated RNA with a log2FC of greater than 2.0. Exp=expression, P-val=P-value

4.2.3ii Downregulated genes in nodular BCC

Genes found to be significantly downregulated were organised according to their Log2FC (Table 4.8) and then according to the highest expressed (Table 4.9). Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Exp	Log2FC	P val	Gene	Exp	Log2FC	P val
SLC39A2	2.4	-4.6	<0.001	HDC	2.6	-3.6	0.002
DGKQ	2.1	-3.8	0.005	FAM65C	2.6	-3.5	<0.001
KRT6A	7.9	-3.7	0.003	GBP6	2.2	-3.5	0.009
CYTH4	2.9	-3.6	0.006	GNLY	2.9	-3.5	0.003
CD2	4.5	-3.6	0.002	CD1C	2.4	-3.5	0.005

Table 4.8 Log2FC downregulated genes in nodBCC. Top 10 differentially expressed RNA are arranged according to the lowest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Exp	Log2FC	P val	Gene	Exp	Log2FC	P val
KRT6A	7.9	-3.7	0.003	KRT16	6.5	-2.9	0.004
FABP5	7.9	-3.4	0.001	LYPD3	6.4	-2.8	0.004
KRT6B	7.1	-2.1	0.006	CCL5	6.4	-2.8	0.005
SERPINB1	6.8	-2.2	0.005	LAD1	5.3	-2.2	0.003
CCL21	6.7	-2.4	0.005	IGHG4	5.1	-2.5	<0.001

Table 4.9 Most downregulated genes in nodBCC arranged according to expression levels. Top 10 differentially, highest expressed, downregulated RNA with a log2FC of less than -2.0. Exp=expression, P-val=P-value.

4.2.4 Comparison of BCC subtypes

A total of 128 differentially expressed genes were seen when comparing tumour subtypes only. The heatmap below (Fig 4.4) gives an overview of the over and under expressed genes seen within each subtype when compared against each other. Log2FC ranged from -6.1 to 8.0 with the majority being up regulated in morphoeic tumour and only 31 genes up regulated in nodular. Expression levels ranged from -2.2 to 7.4.

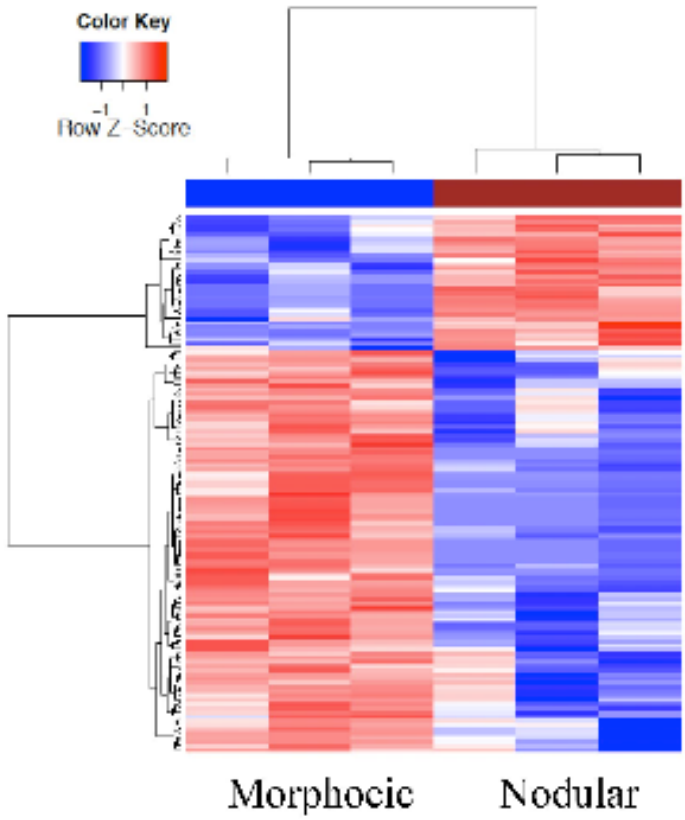


Fig 4.4 Heatmap comparing mBCC and nodBCC tumour. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression.

4.2.4i Upregulated in morphoeic compared to nodular BCC

Genes found to be significantly upregulated were organised according to their Log2FC (Table 4.10) and then according to the highest expressed (Table 4.11). Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
KRT77	2.8	5.9	0.007	ZNF697	2.7	4.0	0.004
OASL	2.1	4.9	0.006	ITGA11	3.7	3.7	0.004
DGKQ	2.1	4.7	0.001	PKP1	7.4	3.1	0.002
CD1C	2.4	4.5	0.003	GPC1	4.5	3.0	0.010
NAB2	3.2	4.0	0.004	DOK4	3.0	2.9	0.004

Table 4.10 Highest Log2FC upregulated genes in mBCC compared to nodBCC. Top 10 differentially expressed RNA are arranged according to the highest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
PKP1	7.4	3.1	0.002	USP5	5.1	2.6	0.004
APCDD1	7.0	2.0	0.001	UNC5B	4.8	2.8	0.007
LAMB3	6.4	2.3	0.006	SLC38A5	4.6	2.4	0.008
WDR46	5.2	2.5	0.006	ACTR1B	5.6	2.5	0.010
EPHB3	5.2	2.7	0.003	GPC1	4.5	3.0	0.007

Table 4.11 Most upregulated genes in mBCC compared to nodBCC and arranged according to expression levels. Top 10 differentially, highest expressed, upregulated RNA with a log2FC of greater than 2.0. Exp=expression, P-val=P-value.

4.2.4ii Downregulated genes in morphoeic compared to nodular BCC

Genes found to be significantly downregulated were organised according to their Log2FC (Table 4.12) and then according to the highest expressed (Table 4.13).

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
Hsa-miR-3117	2.2	-3.7	0.003	CBWD5	4.0	-2.1	0.003
SAMD12	3.9	-2.8	0.004	NCAM1	4.5	-1.9	0.009
NT5DC3	4.5	-2.7	0.002	TMTC1	5.7	-1.6	0.003
MAN1C1	4.0	-2.5	0.006	TJP1	6.6	-1.6	<0.001
TNFAIP2	6.6	-2.4	<0.001	RND3	6.6	-1.4	0.005

Table 4.12 Log2FC downregulated genes in mBCC compared to nodBCC. Top 10 differentially expressed RNA are arranged according to the lowest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
Hsa-miR-3117	2.2	-3.7	0.003	TNFAIP2	6.6	-2.4	<0.001
SAMD12	3.9	-2.8	0.004	NCAM1	4.5	-1.9	
MAN1C1	4.0	-2.5	0.006	TMTC1	5.7	-1.6	
CBWD5	4.0	-2.1	0.003	TJP1	6.6	-1.6	
NT5DC3	4.5	-2.7	0.002	RND3	6.6	-1.4	

Table 4.13 Most downregulated genes in mBCC compared to nodBCC and arranged according to expression levels. Top 6 differentially, highest expressed, downregulated RNA with a log2FC of less than -2.0. The last four are not statistically significant. Exp=expression, P-val=P-value.

4.2.4iii Shared genes present in both morphoeic and nodular tumour

Further analysis was done to take into expression across both tumour subtypes and their corresponding normal tissue (Figure 4.5). The shared genes between the two numbered 44 and were plotted to reveal closer concordance with the upregulated genes (Figure 4.6). Explanations of a selected shared gene are mentioned below noting their biological relevance to cancer.

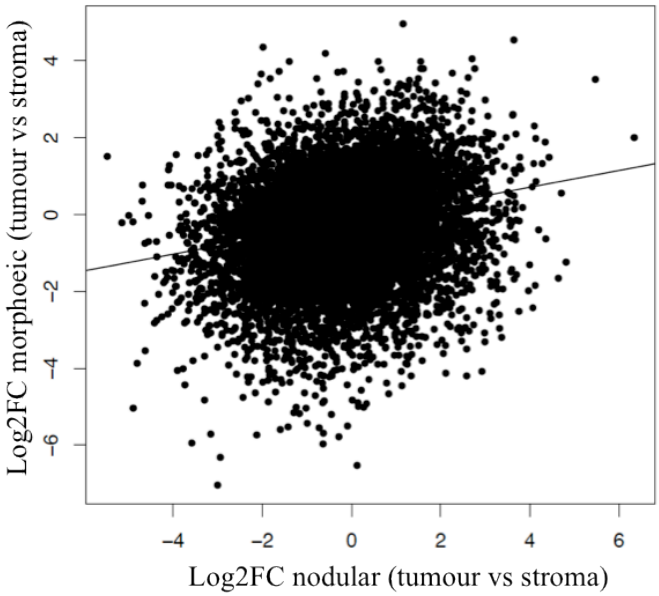


Fig 4.5 RNA expression between morphoeic and nodular BCC subtypes. All filtered genes that were tested are represented as log2FC expression in the above graph with a line that best fits. Selection of the 44 shared genes are shown below in Figure 4.6.

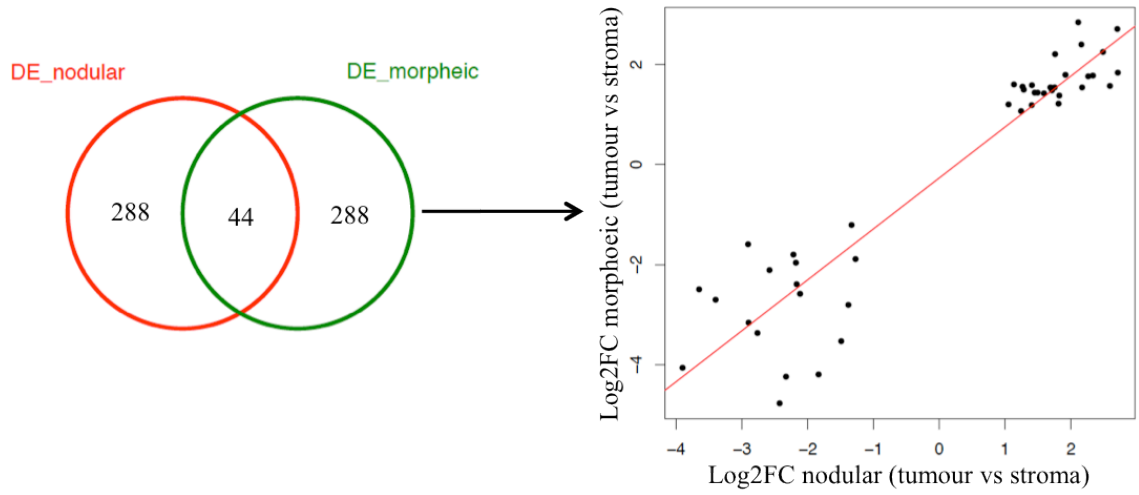


Fig 4.6 Shared differentially expressed genes across both morphoeic and nodular BCC subtypes. The Venn diagram shows each subtype with 288 differentially expressed (DE) genes and 44 were shared between the two. These genes are then represented as a Log2FC expression with a line that best fits. The genes that are upregulated show greater concordance than the downregulated genes,

4.2.5 Morphoeic normal stroma RNA expression

Following from the discovery of DDR1 expression, the normal stromal tissue from both subtypes were compared against each other to see if the mBCC normal stroma had any abnormally expressed genes. This raises the possibility that the surrounding tissue of the mBCC tumour is not entirely normal and as suggested by the WES stroma findings (see section 3.2.9). In total, there are 37 differentially expressed genes seen in the morphoeic stroma compared to the nodular stroma using $p < 0.01$ and log2FC > or <1. The heatmap highlights that all but three genes were upregulated in morphoeic stroma (Figure 4.7).

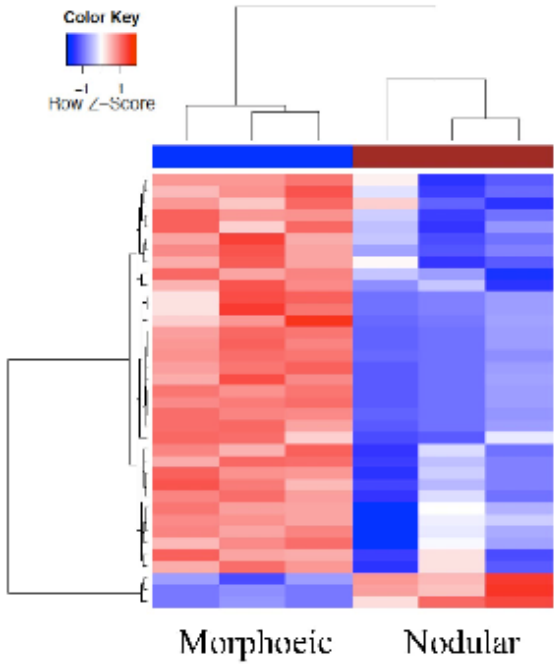


Fig 4.7 Comparison of stromal expression mBCC and nodBCC. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression.

4.2.5i Upregulated genes in morphoeic stroma

DDR1 was discussed above a gene present in both the morphoeic tumour and normal stroma. Comparing morphoeic tumour with nodular normal stroma, the expression was 7.2, log2FC 1.8 and P 0.0005. When comparing the morphoeic stroma with nodular stroma the expression was 7.2, log2FC 0.9 and P = 0.04. Genes found to be significantly upregulated were organised according to their Log2FC (Table 4.14) and then according to the highest expressed (Table 4.15).

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
GSTM1	1.2	7.6	<0.001	CCR2	2.1	3.5	0.005
ACVR1C	0.5	4.9	0.022	SIM2	0.9	3.4	0.008
AC009237.8	2.0	4.6	0.012	ZNF446	1.1	3.1	0.009
CD1B	0.5	4.2	0.007	LPXN	4.1	2.1	0.007
RSPH4A	0.9	4.0	0.008	ABHD5	4.3	1.5	0.007

Table 4.14 Highest Log2FC upregulated genes in the stroma of mBCC compared to nodular stroma. Top 10 differentially expressed RNA are arranged according to the highest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Glutathione S-Transferase Mu 1 (GSTM1) is a detoxification enzyme that is overexpressed in morphoeic stroma suggesting an environment undergoing oxidative stress. Other genes of interest include Activin A Receptor Type 1C (ACVR1C), a TGFB receptor, a pathway that may be BCC pathway and it binds SMAD transcription regulators. CD1b Molecule (CD1B) is an antigen binding protein that triggers T-cell and NK cells. C-C Motif Chemokine Receptor 2 (CCR2) is a chemokine that attracts monocytes, particularly in the inflammatory response to tumours. Single-Minded Family BHLH Transcription Factor 2 (SIM2) is a master gene for CNS development and may either be irrelevant or be related to the axonal guidance that seems to be upregulated in mBCC. Leupaxin (LPXN) contributes to the regulation of cell migration, adhesion and spreading and may make the environment permissible for tumour spread.

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
ABHD5	4.4	1.5	0.007	ZNF446	1.1	3.1	0.009
LPXN	4.1	2.1	0.07	RSPH4A	0.9	4.0	0.008
CCR2	2.1	3.5	0.005	SIM2	0.9	3.4	0.008
ACOO9237.8	2.0	4.6	0.001	CD1B	0.5	4.2	0.007
GSTM1	1.2	7.6	<0.001	ACVR1C	0.5	4.9	0.002

Table 4.15 Most downregulated genes in mBCC stroma compared to nodBCC stroma and arranged according to expression levels. Top 10 differentially, highest expressed, upregulated RNA with a log2FC of greater than 2.0. Exp=expression, P-val=P-value.

4.2.5ii Downregulated genes in morphoeic stroma

As mentioned, only three genes were downregulated and they had minimal expression. DMRT Like Family A1 (DMRTA1) is probably involved in transcription and may be associated with type 2 DM. Little is known about the other two.

Gene	Expression	Log FC	P value
DMRTA1	0.4	-5.2	0.005
RP13-15M17.1	1.8	-4.1	0.008
CTD-2257D19.6	1.5	-3.5	0.007

Table 4.16 Log2FC downregulated genes in the stroma mBCC. Top 3 differentially expressed RNA are arranged according to the lowest log2FC with an average expression greater than 2.0. No other genes were statistically significant. Exp=expression, P-val=P-value

4.2.6 RNA pathway analysis

RNA sequencing expression was assessed using GSEA to identify significantly altered genesets. Pathway expression was first assessed on an individual tumour basis by comparing tumour with the stroma followed by a comparison between the two subtypes. Significant pathway alteration was determined using a FDR q-value ≤ 0.10. Following, subtypes are compared at an individual gene level for selected pathways of interest (see Appendix 1).

4.2.6i Morphoeic BCC expression

Eleven pathways were shown to be upregulated when analysing RNAseq gene-set data (Table 4.17) The ribosome pathway came out as the most upregulated, with BCC and Hh pathway close behind. The TGF beta pathway is also significantly involved and may act in concert with the other pathways. More pathways were downregulated, with 34 significantly reduced (Table 4.18) A lot of these are immune pathways such as the NK, antigen presenting, complement cascades and Fc epsilon, suggesting that a localised immunodeficiency assists local mBCC behaviour.

MSigDB pathway	NES	Geneset size	q-val	P-val
Ribosome	2.65	85	<0.001	<0.001
Basal_cell_carcinoma	2.30	46	<0.001	<0.001
Hedgehog_signaling_pathway	2.17	42	8.5E-04	<0.001
Spliceosome	1.92	125	0.009	<0.001
Base_excision_repair	1.91	31	0.009	<0.001
Dna_replication	1.69	35	0.062	0.003
Basal_transcription_factors	1.67	31	0.064	<0.001
Arrhythmogenic_right_ventricular_cardiomyopathy_arvc	1.62	67	0.088	<0.001
Cysteine_and_methionine_metabolism	1.62	29	0.079	0.016
Tgf_beta_signaling_pathway	1.58	75	0.099	<0.001

Table 4.17 Upregulated mBCC pathways using GSEA. Upregulated MSigDB pathway expression in morphoeic tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.10. MSigDB, molecular signature database; NES, normalised enrichment score; Q-value, false discovery rate; P-val, P-value.

MSigDB pathway	NES	GS size	q-val	P-val
Allograft_rejection	-2.38	29	<0.001	<0.001
Graft_versus_host_disease	-2.36	30	<0.001	<0.001
Hematopoietic_cell_lineage	-2.23	67	<0.001	<0.001
Autoimmune_thyroid_disease	-2.22	31	<0.001	<0.001
Antigen_processing_and_presentation	-2.20	56	<0.001	<0.001
Primary_immunodeficiency	-2.18	28	<0.001	<0.001
Natural_killer_cell_mediated_cytotoxicity	-2.13	98	1.2E-04	<0.001
Complement_and_coagulation_cascades	-2.13	46	1.0E-04	0.002
Asthma	-2.10	21	9.2E-05	<0.001
Type_1_diabetes_mellitus	-2.10	33	8.3E-05	<0.001
Intestinal_immune_network_for_iga_production	-2.07	35	1.5E-04	<0.001
Cytokine_cytokine_receptor_interaction	-2.03	186	2.8E-04	<0.001
Leishmania_infection	-2.02	62	3.9E-04	<0.001
Chemokine_signaling_pathway	-1.90	159	0.001	<0.001
Cell_adhesion_molecules_cams	-1.89	116	0.002	<0.001
T_cell_receptor_signaling_pathway	-1.89	99	0.002	<0.001
Abc_transporters	-1.86	40	0.003	<0.001
Glycosphingolipid_biosynthesis_lacto_and_neolacto_series	-1.83	22	0.004	0.003
B_cell_receptor_signaling_pathway	-1.79	70	0.006	<0.001
Viral_myocarditis	-1.72	67	0.014	0.002
Rig_1_like_receptor_signaling_pathway	-1.71	51	0.014	0.003
Arachidonic_acid_metabolism	-1.62	44	0.031	0.008
Systemic_lupus_erythematosus	-1.60	69	0.039	0.004
Proximal_tubule_bicarbonate_reclamation	-1.59	15	0.040	0.030
Toll_like_receptor_signaling_pathway	-1.59	81	0.039	0.004
O_glycan_biosynthesis	-1.59	22	0.039	0.025
Neuroactive_ligand_receptor_interaction	-1.58	126	0.040	0.001
Nod_like_receptor_signaling_pathway	-1.56	59	0.046	0.011
Cytosolic_dna_sensing_pathway	-1.55	40	0.050	0.021
Retinol_metabolism	-1.53	34	0.056	0.033
Leukocyte_transendothelial_migration	-1.52	101	0.062	0.018
Fc_epsilon_ri_signaling_pathway	-1.50	66	0.073	0.020
Apoptosis	-1.48	76	0.082	0.021
Phosphatidylinositol_signaling_system	-1.45	71	0.098	0.023
Allograft_rejection	-2.38	29	<0.001	<0.001
Graft_versus_host_disease	-2.36	30	<0.001	<0.001
Hematopoietic_cell_lineage	-2.23	67	<0.001	<0.001

Table 4.18 Downregulated mBCC pathways using GSEA. Downregulated MSigDB pathway expression in morphoeic tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.10 MSigDB, molecular signature database; NES, normalised enrichment score; Q-value, false discovery rate; P-val, P-value MSigDB, molecular signature database; NES, normalised enrichment score; GS size, gene-set size; FDR false discovery rate

4.2.6ii Nodular BCC

Only two pathways are upregulated significantly in nodBCC when assessing RNAseq GSEA (Table 4.19). BCC pathway is upregulated, but most notably the Hh pathway does not feature. In contrast, 40 pathways are down regulated in nodBCC including immune pathways such as NK, Toll like receptor, complement cascades, but there are many others such as type 1 diabetes mellitus, asthma and Parkinson's disease whose biological significant seems unrelated (Table 4.20).

MSigDB pathway	NES	Geneset size	q-val	P-val
Mismatch_repair	1.84	22	0.069	0.002
Basal_cell_carcinoma	1.82	46	0.041	<0.001

Table 4.19 Upregulated nodBCC pathways using GSEA. Upregulated MSigDB pathway expression in morphoeic tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.10. MSigDB, molecular signature database; NES, normalised enrichment score; Q-value, false discovery rate; P-val, P-value.

MSigDB pathway	NES	GS size	q-val	Pval
Ribosome	-2.86	85	<0.001	<0.001
Hematopoietic_cell_lineage	-2.63	67	<0.001	<0.001
Antigen_processing_and_presentation	-2.61	56	<0.001	<0.001
Graft_versus_host_disease	-2.57	30	<0.001	<0.001
Allograft_rejection	-2.46	29	<0.001	<0.001
Type_1_diabetes_mellitus	-2.43	33	<0.001	<0.001
Oxidative_phosphorylation	-2.41	108	<0.001	<0.001
Cytokine_cytokine_receptor_interaction	-2.33	186	<0.001	<0.001
Autoimmune_thyroid_disease	-2.32	31	<0.001	<0.001
Systemic_lupus_erythematosus	-2.24	69	<0.001	<0.001
Intestinal_immune_network_for_iga_production	-2.22	35	<0.001	<0.001
Asthma	-2.21	21	<0.001	<0.001
Parkinsons_disease	-2.13	105	1.7E-04	<0.001
Primary_immunodeficiency	-2.13	28	1.6E-04	<0.001
Leishmania_infection	-2.13	62	1.5E-04	<0.001
Toll_like_receptor_signaling_pathway	-2.04	81	6.9E-04	<0.001
Cell_adhesion_molecules_cams	-2.01	116	9.4E-04	<0.001
Natural_killer_cell_mediated_cytotoxicity	-1.96	98	0.002	<0.001
Complement_and_coagulation_cascades	-1.95	46	0.002	<0.001
Alzheimers_disease	-1.95	147	0.002	<0.001
T_cell_receptor_signaling_pathway	-1.94	99	0.002	<0.001
Nod_like_receptor_signaling_pathway	-1.92	59	0.002	<0.001
Proteasome	-1.88	42	0.003	0.002
Huntingtons_disease	-1.87	158	0.003	<0.001
Prion_diseases	-1.82	29	0.005	0.004
Chemokine_signaling_pathway	-1.81	159	0.006	<0.001
Pantothenate_and_coa_biosynthesis	-1.77	15	0.009	0.004
B_cell_receptor_signaling_pathway	-1.71	70	0.016	0.002
Cardiac_muscle_contraction	-1.69	61	0.019	0.008
Viral_myocarditis	-1.67	67	0.021	0.004
Jak_stat_signaling_pathway	-1.67	111	0.021	<0.001
Snare_interactions_in_vesicular_transport	-1.61	38	0.034	0.006
Tyrosine_metabolism	-1.61	32	0.033	0.008
Cytosolic_dna_sensing_pathway	-1.56	40	0.049	0.022
Arachidonic_acid_metabolism	-1.54	44	0.054	0.026
PPAR_signaling_pathway	-1.54	51	0.053	0.011
Rig_1_like_receptor_signaling_pathway	-1.51	51	0.063	0.024

Table 4.20 Downregulated nodBCC pathways using GSEA. Downregulated MSigDB pathway expression in morphoeic tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.10 MSigDB, molecular signature database; NES, normalised enrichment score; Q-value, false discovery rate; P-val, P-value MSigDB, molecular signature database; NES, normalised enrichment score; GS size, geneset size; FDR false discovery rate.

4.2.6iii Morphoeic versus Nodular tumour

The two tumour subtypes were then compared in terms of pathway analysis and it demonstrates a shared expression of BCC and several immune pathways (Figure 4.8) Pathways that were more expressed in morphoeic than nodular are shown in Table 4.21 and include extracellular matrix interaction and cell adhesion; which fits mBCC phenotype. In addition, notch and axonal guidance signalling seem to play a role in mBCC. Individual gene level expression is shown for selected pathways in Appendix 1.

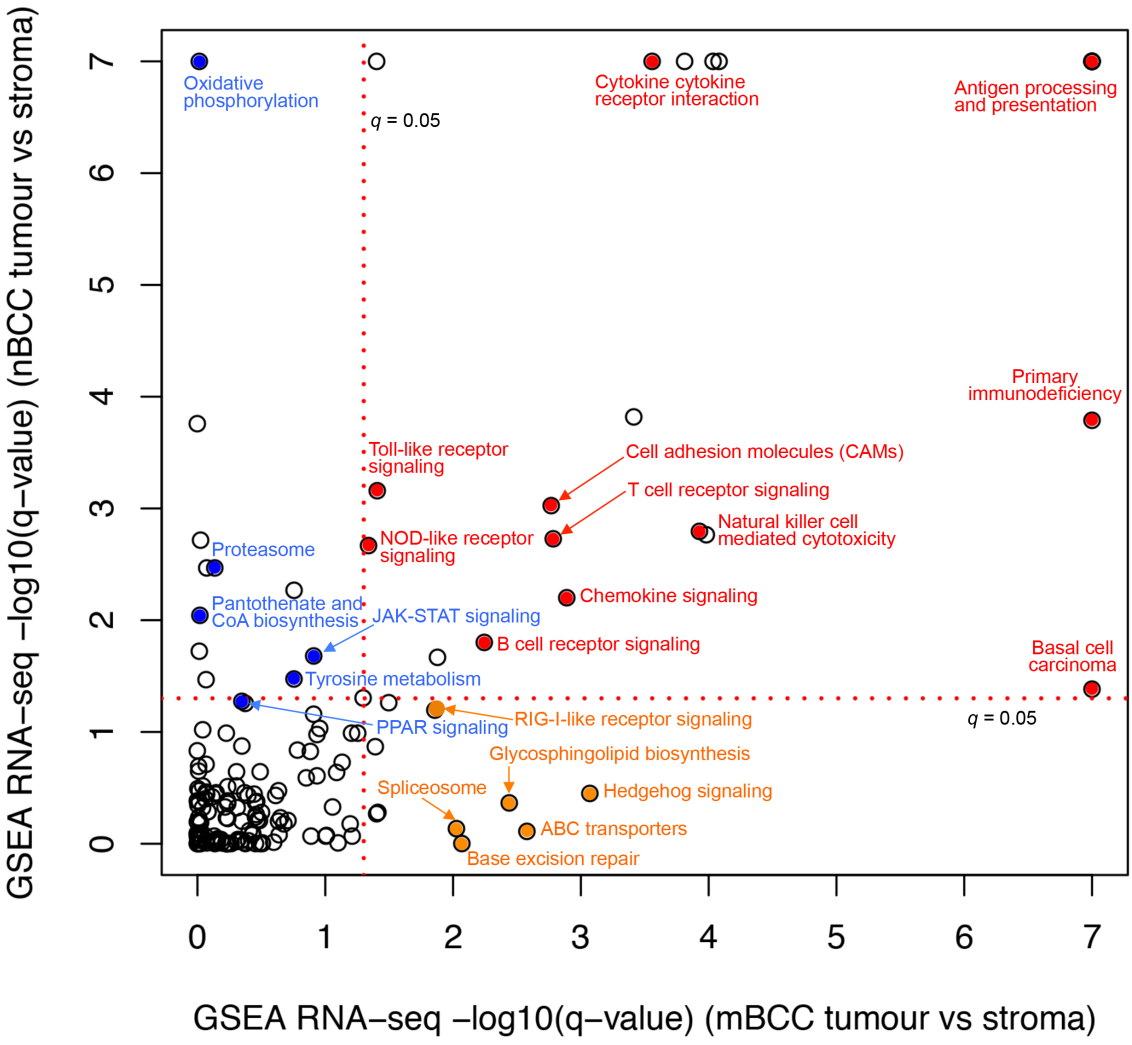


Figure 4.8 Pathway comparison between morphoeic and nodular BCC. Pathways identified using gene-set enrichment analysis (GSEA) of RNAseq data for morphoeic and nodular BCC are represented. A selection of shared pathways are highlighted in red. Blue represents those significantly altered pathways in nodular BCC and those pathways in orange represent morphoeic specific pathways. nodBCC, nodular BCC; GSEA, gene-set enrichment analysis; mBCC, morphoeic BCC; q, q-value.

MSigDB pathway	NES	Geneset Size	q-val	P-val
Ribosome	2.58	85	<0.001	<0.001
Ecm_receptor_interaction	1.98	76	0.003	<0.001
Systemic_lupus_erythematosus	1.98	69	0.002	<0.001
Intestinal_immune_network_for_iga_production	1.82	35	0.024	<0.001
Antigen_processing_and_presentation	1.81	56	0.019	0.0012
Asthma	1.75	21	0.036	0.004
Type_1_diabetes_mellitus	1.73	33	0.042	0.007
Vasopressin_regulated_water_reabsorption	1.72	38	0.041	0.004
Allograft_rejection	1.66	29	0.072	0.009
Cell_adhesion_molecules_cams	1.63	116	0.087	<0.001
Hematopoietic_cell_lineage	1.63	67	0.080	0.001
Autoimmune_thyroid_disease	1.60	31	0.105	0.016
Notch_signaling_pathway	1.59	43	0.103	0.015
Graft_versus_host_disease	1.58	30	0.105	0.017
Pantothenate_and_coa_biosynthesis	1.58	15	0.103	0.033
Axon_guidance	1.56	115	0.107	0.002

Table 4.21 Upregulated mBCC pathways compared to nodBCC using GSEA. MSigDB Pathway expression when comparing morphoeic tumour versus nodular tumour. All pathways are upregulated in morphoeic tumour and arranged according to NES with a false discovery rate (q-value) <0.10. MSigDB, molecular signature database; NES; FDR false discovery rate.

4.2.7 RNA sequencing validation

Validation of two genes (PDGFA and DDR1) was performed using SYBR RT-qPCR. RNA from 3 morphoeic and 3 nodular samples were analysed against normal eyelid tissue from a non-cancerous eye. PDGFA was found to be one of the 44 shared genes in the RNAseq analysis and may contribute to extracellular matrix modification of the tumour. On the other hand, DDR1 expression using RNAseq was only found in the morphoeic tumour and stroma.

Relative expression levels (x20) of PDGFA demonstrated increased expression in both mBCC and nodBCC when compared to normal eyelid tissue (Fig4.9). There was a trend to increased expression in mBCC compared to nodBCC, however this was non-significant. As a potent mitogen of mesenchymal cells this confirms the shared gene expression on RNAseq data and suggests a significant role in both subtypes.

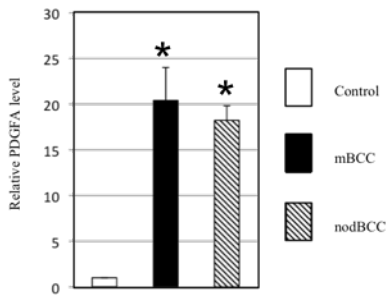


Fig 4.9 PDGFA q-rtPCR expression in mBCC and nodBCC. Relative PDGFA levels in both mBCC and nod BCC compared to normal eyelid whilst each were normalised to housekeeping gene B-actin. N=3 for each sample type. *P<0.01 compared to normal eyelid

In contrast, DDR1 levels were shown to be less than normal eyelid tissue for both subtypes (Fig 4.10). Interestingly, there is a slight difference between levels in the mBCC and nodBCC, being more in the former, which is what the RNAseq data suggested, however, the fact that the expression is less than normal eyelid implies it may not be important. Nevertheless, the RNAseq data highlighted its presence in the mBCC stroma and the normal eyelid sample contains tarsal plate, skin and stroma, thus it may be better to compare normal stroma with mBCC stroma in the future to see if it has a significant role in this microenvironment.

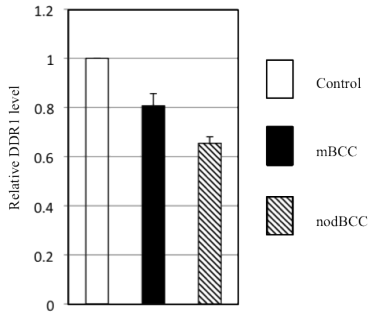


Fig 4.10 DDR1 q-rtPCR expression in mBCC and nodBCC. Relative DDR1 levels in both mBCC and nod BCC compared to normal eyelid whilst each were normalised to housekeeping gene B-actin. No P value as no significant difference seen between samples.

4.3 Correlated WES and RNA sequencing data

4.3.1 Individual driver genes with RNA changes in BCC

Driver genes identified using WES data were assessed in terms of their expression. mBCC driver genes (see section 3.2.4) are listed in Table 4.22. Correlation of RNAseq and WES data revealed increased expression for PTCH1 as the only significantly changed using a $P < 0.05$. A trend towards down regulation for HECTD4 and up regulation for SMARCA4 is noted. The reverse trend could be expected if HECTD4 is involved in tumorigenesis and SMARCA4 acts as a tumour suppressor gene. However, the log2FC was rather low so on an individual gene basis it is not clear what could be happening on a functional level. PTCH1 is upregulated presumably in an attempt to suppress the overactive Hh pathway, so it may be better to look at an overall pathway picture rather than an individual gene level change for both SMARCA4 and HECTD4. Loss of function in the gene cannot be ruled out as a cause of low expression or log2FC.

Driver	Exp	Log2FC	P-val	Driver	Exp	Log2FC	P-val
PTCH1	5.7	1.2	0.01	SMARCA4	5.6	0.50	0.10
FLNB	7.3	-0.16	0.63	FNIP1	6.1	-0.1	0.85
TTN	8.3	-0.58	0.21	CHKB	-	-	-
USP4	6.6	-0.04	0.91	CCDC108	-	-	-
CHD3	5.8	0.09	0.81	FBXO15	0.8	0.3	0.6
EPHA3	1.8	-0.50	0.64	POM121L12	3.3	-0.3	0.7
NFS1	3.9	0.03	0.96	OR5H2	-	-	-
EFTUD2	6.5	-0.10	0.81	KRTAP12-3	-1.0	2.0	0.24
HECTD4	5.1	-0.09	0.09				

Table 4.22 Driver gene expression in mBCC. WES drivers selected using Intogen and MutSigCV with a $q<0.02$ and $p<0.02$ respectively.

Driver genes identified using WES for nodBCC (see section 3.2.6) were assessed for their RNA expression in Table 4.23. Correlation of WES and RNAseq data revealed two genes, PTCH1 and COL4A4. Increased expression in PTCH1 is seen as in mBCC presumably for the same reasons. COL4A4 is greatly reduced in expression suggesting a loss of function. Loss of the 5-prime end of COL4A4 results in the development of leiomyomatoma which behave in an indolent fashion rather like nodBCC and full loss of function is seen in oesophageal cancer.

Driver	Exp	Log2FC	P-val	Driver	Exp	Log2FC	P-val
PTCH1	5.7	1.8	0.002	MTR	6.2	-0.3	0.388
TP53	7.0	0.3	0.377	MYCN	4.4	0.8	0.182
FADS1	5.3	-0.6	0.091	NARS	7.4	-0.3	0.921
PPM1D	3.2	-0.1	0.861	GABRA6	-	-	-
TAF4B	3.0	-0.3	0.72	TSGA13	0.8	0.8	0.528
TRIP4	5.1	0.1	0.846	OR52J3	-	-	-
COL4A4	2.0	-2.0	0.033	TRIM39	4.0	0.8	0.248
ARGAP35	-	-	-	C10orf82	1.0	0.7	0.463

Table 4.23 Driver gene expression in nodBCC. WES drivers selected using Intogen and MutSigCV with a $q<0.02$ and $p<0.02$ respectively.

Potentially important genes within the morphoeic stroma were assessed at a RNA expression level, however, none were significantly expressed (Table 4.24).

Gene	Exp	Log2FC	P-val	Gene	Exp	Log2FC	P-val
PTPN22	2.2	1.4	0.13	*MYBPC1*	1.7	1.3	0.70
ARHGEF28	4.2	1.0	0.19	*XPO4*	4.8	0.1	0.78
ASAP1	6.2	0.16	0.73	*CPNE7*	1.0	2.8	0.07
C5	2.4	1.0	0.33	*CDH20*	0.2	0.6	0.73
SFMBT2	2.0	-0.52	0.62				

Table 4.24 Correlation of stromal mBCC WES and RNAseq pathway analysis. Top ten loss/gain mutations seen in morphoeic BCC stroma compared to RNA expression of morph stroma.

4.3.2 Pathway link of DNA and RNA analysis

In a similar fashion to the driver genes, significantly modified pathways detected using WES data was compared to pathways picked up using RNAseq. Comparison for mBCC data is summarised in Figure 4.11 and Table 4.25. Upregulation of Hh signalling is significant for both WES and RNAseq analysis and loss of function in NK and FC epsilon pathways seem important. Axonal guidance is not altered in terms of WES analysis, but significantly overexpressed using RNAseq pathway detection (Figure 4.12). As mentioned, it may work in concert with Wnt and Hh in mBCC tumorigenesis.

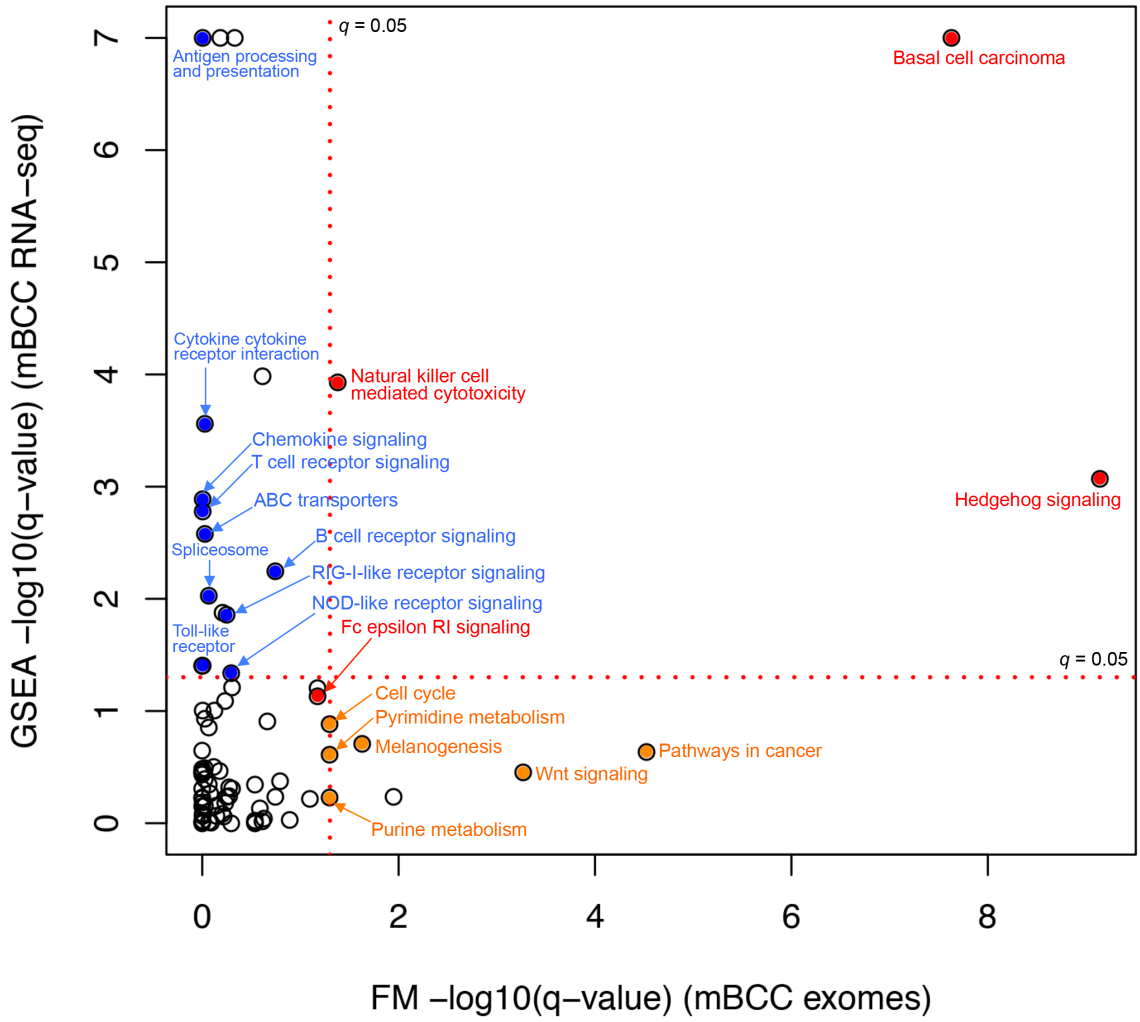


Figure 4.11 Pathways in morphoeic BCC. WES and RNAseq pathway data are compared. Red pathways are significantly changed in DNA and RNA analysis using a q=0.05. Orange pathways represent pathways altered in WES data only and blue pathways are RNAseq expressed pathways; mBCC, morphoeic BCC; GSEA, gene-set enrichment analysis; q, q-value.

	WES			RNAseq			
MSigDB pathway	**Kegg ID**	**q-val**	**F**	**NES**	**GS size**	**q-val**	**P-val**
Hedgehog_signaling_pathway	Hsa04340	7.23E-10	9	2.17	42	8.5E-04	<0.001
Basal_cell_carcinoma	Hsa05217	2.33E-08	9	2.30	46	<0.001	<0.001
Axon_guidance	Hsa04360	0.53	8	1.56	115	0.107	0.002
Natural_killer_cell_mediated_cytotoxicity	Hsa04650	0.042	8	-2.13	98	1.2E-04	<0.001
Fc_epsilon_ri_signaling_pathway	Hsa04664	0.07	6	-1.50	66	0.073	0.020

Table 4.25 Correlation of mBCC WES and RNAseq pathway analysis. Intogen pathway prediction highlighting altered pathways using WES data from 10 mBCC tumours. MSigDB Pathway expression when comparing morphoeic tumour versus nodular tumour. MSigDB, molecular signature database; Kegg ID, KEGG pathway identification number; q-val, q-value; F, frequency of mutations within tumour sample; NES, normalised enrichment score; GS, geneset size; p-val, P-value.

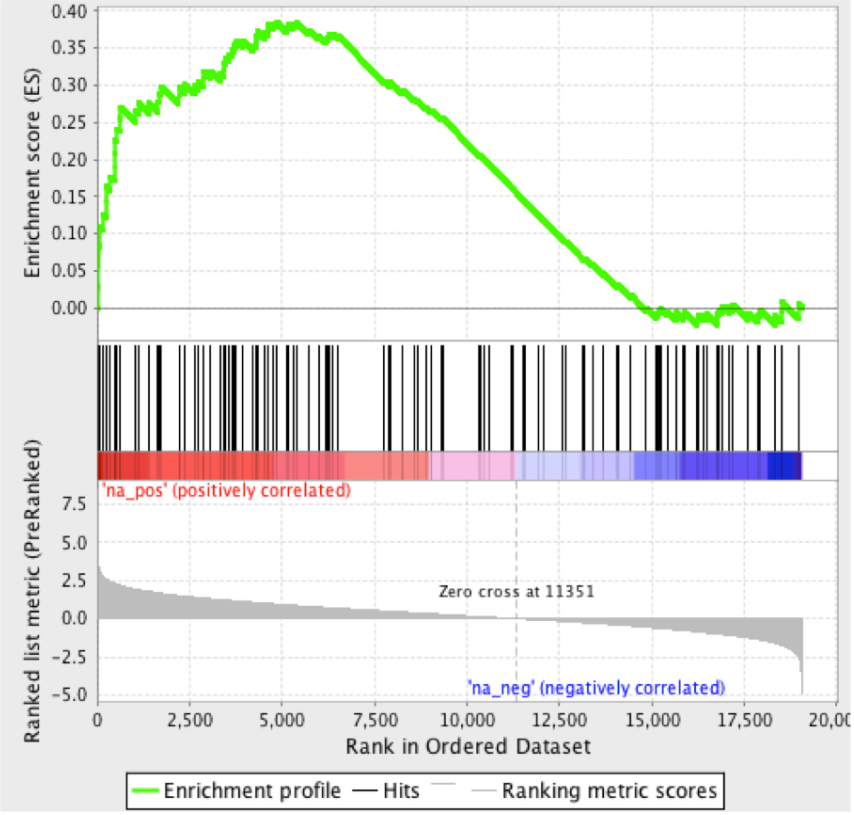


Figure 4.12 Axonal guidance pathway in mBCC. Overall enrichment plot for the 115 genes within the axonal guidance pathway demonstrating an increased normalised enrichment score compared to nodBCC. Each black line represents one of the 115 genes with those to the left most correlated with the pathway.

Nodular BCC revealed three pathways expressed in both WES and RNAseq pathway analysis (Figure 4.13). Hh pathway seems only to be altered at a mutational level and not overexpressed on a RNA level. Overexpression of BCC pathway and loss of PPAR highlights nodBCC activity (Table 4.26).

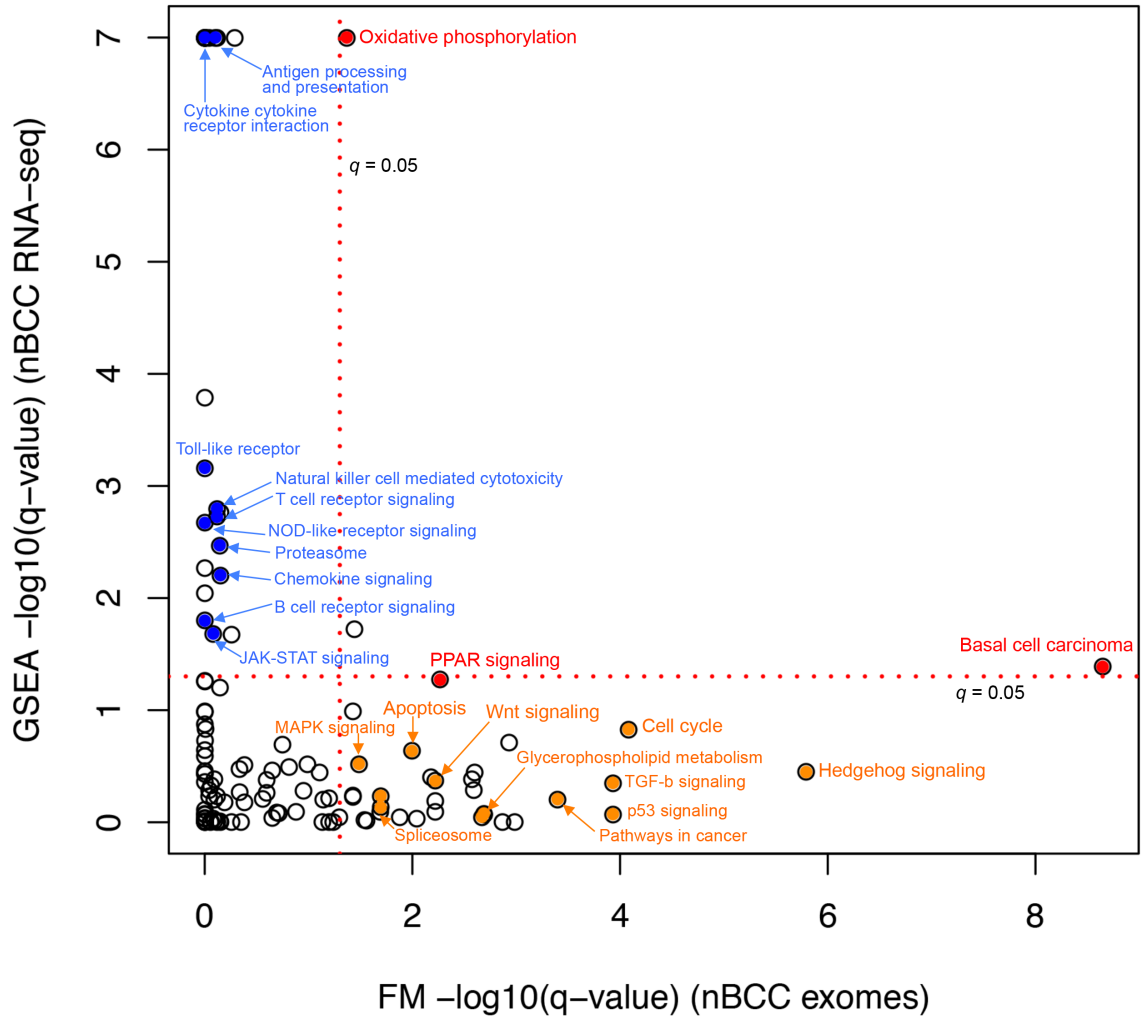


Figure 4.13 Pathways in nodular BCC. WES and RNAseq pathway data are compared. Red pathways are significantly changed in DNA and RNA analysis using a q=0.05. Orange pathways represent pathways altered in WES data only and blue pathways are RNAseq expressed pathways. nodBCC, nodular BCC; GSEA, gene-set enrichment analysis; q, MutsigDB q-value.

	WES			RNAseq			
MSigDB pathway	**ID**	**Q-val**	**F**	**NES**	**GS size**	**q-val**	**P-val**
Mismatch repair	Hsa03430		5	1.84	22	0.069	0.002
Basal_cell_carcinoma	Hsa05217	2.25E-09	9	1.82	46	0.041	<0.001
TGFB	Hsa04350	1.2E-0.4	9	1.03	75	0.846	0.379
P53	Hsa04115	1.2E-0.4	10	-1.09	65	0.448	0.310
PPAR	hsa03320	0.005	9	-1.54	51	0.053	0.011
Oxidative phosphorylation	Hsa00190	0.04	9	-2.41	108	<0.001	<0.001

Table 4.26 Correlation of nodBCC WES and RNAseq pathway analysis. Intogen pathway prediction highlighting altered pathways using WES data from 10 nodBCC tumours. MSigDB Pathway expression when comparing morphoeic tumour versus nodular tumour. MSigDB, molecular signature database; Kegg ID, KEGG pathway identification number; q-val, q-value; F, frequency of mutations within tumour sample; NES, normalised enrichment score; GS, geneset size; p-val, P-value

4.4 Protein expression in BCC

4.4.1 Clinical features of BCC patients

	Number	Mean Age	Sex (M/F)
mBCC	15	71	(9/6)
nodBCC	15	64	(4/11)

Table 4.27 Clinical details of protein expression study in BCC. nodBCC, nodular BCC; mBCC, morphoeic BCC; M, male; F, female

4.4.2 Hedgehog expression in periocular BCC

Signalling of the Hh pathway is known to be a fundamental component of BCC initiation and growth as blocking of it using SMO inhibitors has shown to cause an impressive reduction in tumour size or temporary resolution. It is not clear if the pathway is different between the subtypes and the WES analysis places it as the top pathway in mBCC and the second most important pathway in nodBCC. GSEA showed Hh to be the third upregulated pathway in mBCC, but was not up regulated in nodBCC. Protein expression was therefore sort to see if the there was a difference in expression at a protein level and to see if the canonical pathway was activated. Expression was noted in all four parts of the canonical Hh pathway for both subtypes although nodBCC expressed less than expected especially in Gli1 (Fig 4.14). Semi-quantification using immunofluorescence revealed an increase in all components of the Hh pathway in mBCC tumour compared to nodBCC tumour and positive control tissue (Fig 4.15A). NodBCC showed a trend to an increase in Gli2 expression but the rest of the pathway was less than the positive control tissue, nevertheless still present. Interestingly, the stroma of mBCC demonstrated Hh expression (Fig 4.15B) and the RNAseq expression data showed a trend towards this finding (Table 4.28).

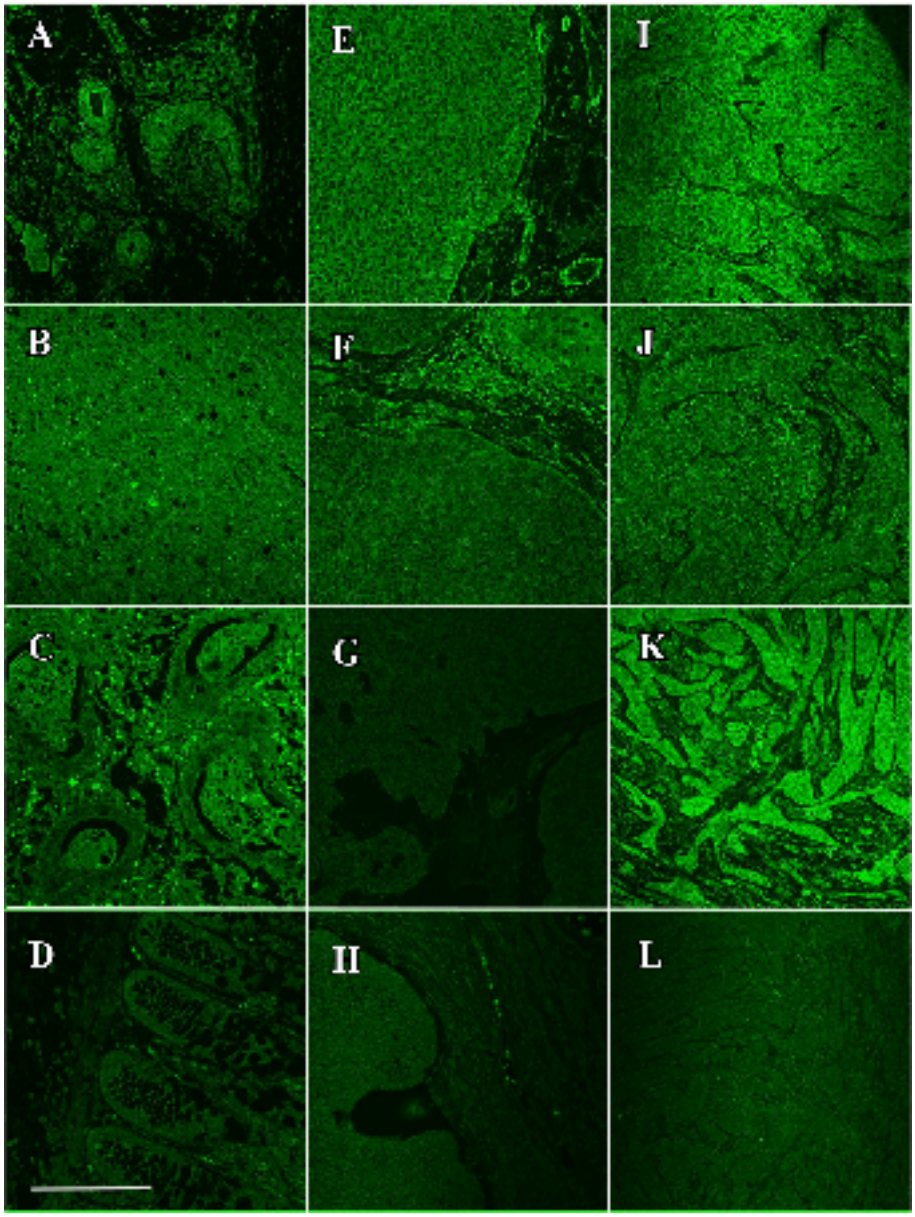
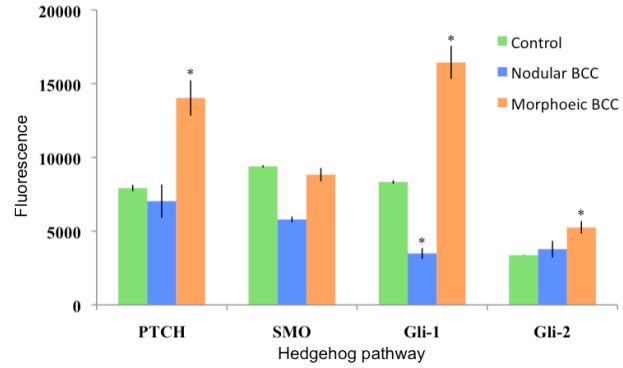


Fig 4.14 Hh protein expression in BCC. Representative pictures of Hedgehog pathway expression proteins using immunofluorescence in positive control tissue (A-D), nodBCC (E-H) and mBCC (I-L). Secondary antibody staining utilized AlexaFluor-568 and the colour converted into green using Image J for pictorial purposes. Scale bar represents 250 µm and all images are at 200 X magnification. Antibody stains for PTCH1 (A,E,I), SMO (B,F,J), Gli1 (C,G,K) and Gli2 (D,H,L). PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

A



B

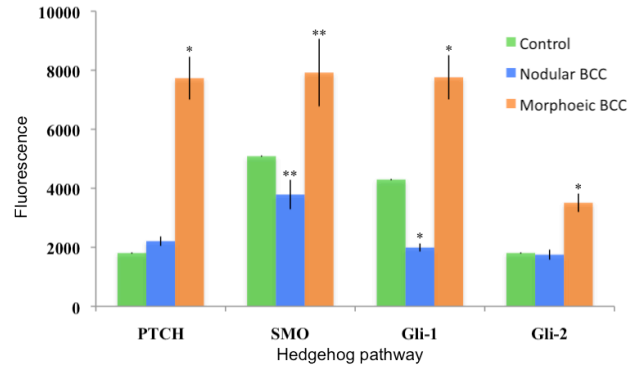


Fig 4.15 Semi-quantification of Hh expression in BCC. Semi-quantification of antibody expression (x-axis) within (A) mBCC tumour, nBCC tumour, control tissue and (B) stroma of mBCC, nBCC, control tissue using fluorescence intensity (y-axis) was determined in regions of interest as delineated by microscopy, ensuring a standardized area size whilst containing the same number of nuclei as determined by DAPI staining. Each tumour sample had an average of three readings. Each bar represents mean values ± SEM taken from 15 nodBCC, 15 mBCC and control tissue samples for each Hh protein. *P<0.01, ** P<0.05 compared to control tissue. PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

RNAseq trend supports protein expression

	Nodular			Morphoeic			M Stroma		
Gene	**Exp**	**Log2FC**	**P-val**	**Exp**	**Log2FC**	**P-val**	**Exp**	**Log2FC**	**P-val**
PTCH1	5.7	1.8	0.002	5.7	3.2	0.04	5.7	2.0	0.20
SMO	4.5	0.2	0.63	4.8	2.0	0.03	4.5	0.8	0.39
Gli1	4.9	1.8	0.006	4.9	3.6	0.02	4.9	1.6	0.34
Gli2	5.1	0.9	0.09	5.1	4.2	0.03	5.1	2.9	0.12

Table 4.28 Hh pathway RNA expression in BCC. RNAseq data supporting protein expression findings of an increased Hh pathway expression in morphoeic tumour compared towards nodular tumour and a trend towards increased stromal expression, especially Gli2.

4.4.3 Glypican 1 expression in periocular BCC

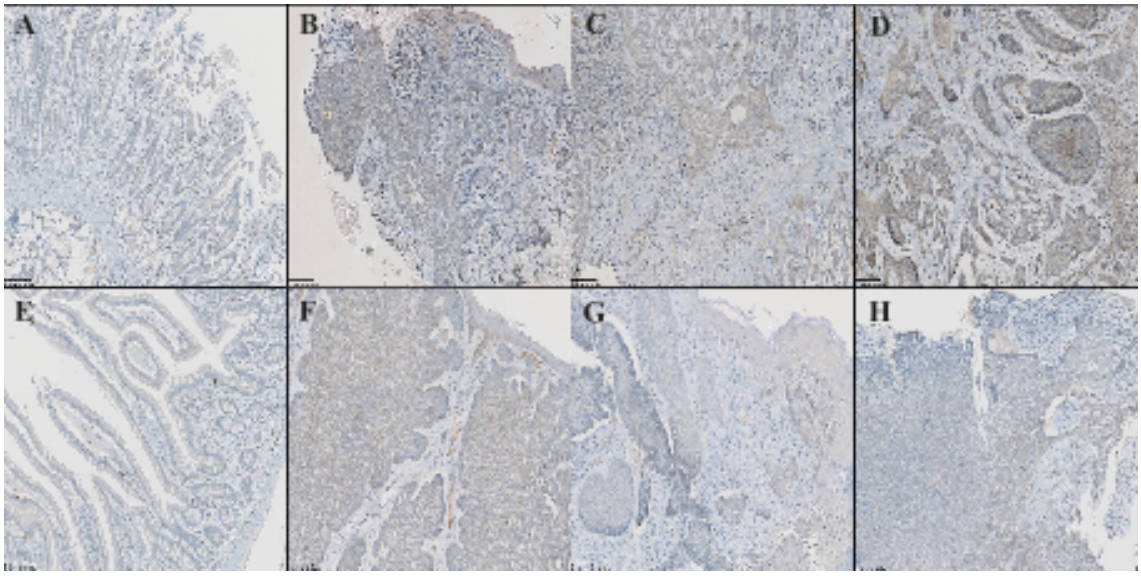


Fig 4.16 Glypican expression in BCC. GPC1 expression in mBCC (B-D) compared to nodBCC (F-H) with duodenum as control tissue (A and E). Scale is 100um. GPC1, Glypican 1.

Glypicans have already been mentioned as important in morphoeic tumours, specifically GPC3 and GPC1. The latter was shown to be unique to mBCC and hence was investigated at a protein level. Increased staining is seen in mBCC tumour compared to nodBCC, although there is some staining in both cancers. GPC1 is thought to inactivate the G1/S checkpoint and stimulate DNA replication with its mitogenic activity described in glioma and pancreatic cancer. Moreover, it acts as a co-receptor in SHH activity and therefore, can potentiate overexpressed Hh activity in mBCC.

4.4.4 Validation of hedgehog pathway antibodies

Western blot of LNCaP-Gli1 cells, a known hyper-expressed Hh pathway cancer cell line, demonstrated good specificity of each antibody with a clear band for the appropriate sized protein (Fig 4.17).

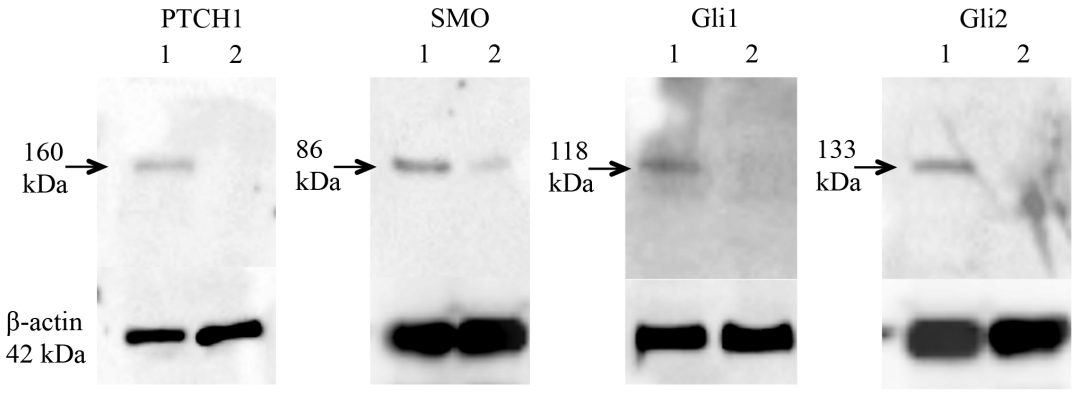


Fig 4.17 Validation of Hh antibodies. Western blot analysis of LnCaP-Gli1 and human wildtype fibroblasts demonstrating good specificity of each antibody with a clear band for the appropriate sized protein. 1=LnCaP-Gli1; 2=Human wildtype fibroblasts; PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

DISCUSSION

4.5 RNA sequencing in periocular BCC

4.5.1 Upregulated genes in morphoeic BCC

Adenomatous Polyposis Coli membrane recruitment protein 2 (AMER2) – chromosome 13q12.13

During embryogenesis, AMER2 is a key negative regulator of the canonical Wnt signalling pathway during neuroectodermal patterning.(Pfister et al., 2012b) Its structure contains binding sites for the tumour suppressor, adenomatous polyposis coli (APC), and interacts with other critical components of the β-catenin destruction complex.(Boutet et al., 2010) Loss of AMER2 induces overexpression of the Wnt/b-catenin pathway. Another function is the recruitment of microtubule-associated protein RP/EB family member (MAPRE1) and knockout of AMER2 reduces cell migration, highlighting its effect on cell migration.(Pfister et al., 2012a) One of the hallmarks of the morphoeic subtype is the widespread cell migration arranged in lines and in skip lesions. It is possible that overexpression of AMER2 contributes to this migratory pattern especially as high expression levels are seen neuron migration during the development of the growing brain.(Comai et al., 2010) Furthermore, excessive AMER2 could be a compensatory response to aberrant Wnt pathway expression.

Anaplastic Lymphoma Receptor Tyrosine Kinase (ALK)- chromosome 2p23.1

This gene is a tyrosine kinase member of the insulin receptor superfamily and has an important function in the genesis and differentiation of the nervous system. The role of ALK in cancer has been extensively documented since it was first found to promote tumour growth by chromosomal rearrangement and fusing to other genes almost 20 years ago.(Ladanyi, 1997) In the case of anaplastic large cell lymphoma there is a nucleophosmin (NPM)/ALK fusion. NPM was highly expressed (expression 10) in morphoeic, but non-significantly up regulated (P=0.27, log2FC 0.4). Fusion of echinoderm microtubule-associated protein-like 4 (EML4)/ALK occurs in a sub group of non-small cell lung cancer (NSCLC) leading to the production of an oncogenic EML4/ALK tyrosine kinase that promotes cell proliferation and is anti-apoptotic.(Rossi et al., 2014) Other oncogenic genetic rearrangements seen in ALK include multiple copies

or a sequence variation that results in constitutive activation. Germline point mutations can occur in ALK causing heritable neuroblastoma.(Mosse et al., 2008) ALK overexpression is seen in breast cancer and related to a more aggressive subtype.(Siraj et al., 2015) ALK overexpression is also seen in inflammatory myofibroblastic tumours and these tumours have some similarity to mBCC under the microscope, namely the florid inflammatory and fibrotic response.(Zhou et al., 2013, Pierry et al., 2015) This data supports the work of Ning et al who demonstrated ALK overexpression in BCC and furthermore, showed that infiltrative (morphoeic) has a higher fold change compared to nodBCC.(Ning et al., 2013) In addition, nodBCC ALK was slightly less upregulated (3.8 expression, 2.1 log2FC and P=0.006) reflecting a similar trend. Moreover, there is likely to be an association with activation of the Hh pathway and Gli1 overexpression.(Singh et al., 2009) There are several ALK inhibitors on the market which might provide an alternative therapeutic modality.(Crescenzo and Inghirami, 2015)

Hydroxysteroid (11-Beta) Dehydrogenase 1 (HSD11B1) – chromosome 1q32.2

This gene codes for an enzyme that interconverts cortisol to cortisone the active/inactive glucocorticoid respectively. Despite being bidirectional, it mainly functions to generate active glucocorticoid and excess is attributed to type 2 diabetes mellitus, obesity, Alzheimer's disease, rheumatoid arthritis and glaucoma.(Longley et al., 2003) Steroid surplus in the skin, as demonstrated by Cushing's syndrome, impairs healing, causes degradation of extracellular matrix proteins and impaired fibroblast function, all of which would confer a benefit for the tumour. A constitutive activating variant (rs932335) in the *HSD11B1* gene poses a risk factor for colorectal cancer in the Chinese population.(Wang et al., 2013a) Overexpression has been shown in colorectal cancer with increasing levels relating to neoplastic transformation. A different variant (rs11807619) is associated with an increased risk of breast cancer. There is therapeutic potential, as several HSD11B1 inhibitors are being developed for type 2 diabetes mellitus and some have been tested with success on human subjects.(Byun et al., 2015, Hong et al., 2015, Wright et al., 2013) Expression was not seen in nodBCC where it was non-significantly downregulated (expression 3.2, log2FC -0.2 and P=0.74) and its presence may confer a more aggressive nature of the tumour.

Peptidyl Arginine deiminase type III (PADI3) – chromosome 1p36.13

Expression of this enzyme that converts arginine into citrulline is restricted to the granular layer of the epidermis and the inner root sheath of the hair follicle.(Dong et al., 2006) Protein citrullination regulates physiological pathways such as gene transcription and in cancer this can be altered. Isoenzyme 4 has been shown to be over expressed in cancer, for example in breast, lung, hepatocellular and ovarian carcinoma.(Chang et al., 2009) Due to the alteration in cancer, rheumatoid arthritis and multiple sclerosis to name a few, the development of PADI3 inhibitors has started.(Knuckley et al., 2010)

Thyroid Stimulating Hormone Receptor (TSHR) – chromosome 14q31.1

TSHR transmits the activity of the thyroid-stimulating hormone (thyrotropin) onto the thyroid gland to release thyroxine and triiodothyronine. Gain of function mutations or hyper expression is seen in thyroid cancer with TSHR mRNA levels acting as a biomarker for malignancy and it can predict metastasis or residual disease.(Chia et al., 2007) Over-expression has been shown to occur in the more aggressive form of breast cancer and may be associated with a poorer prognosis.(Yuan et al., 2014) *TSHR* mutations and overexpression is seen in some cases of lung cancer.(Kim et al., 2012)

SH3 And Cysteine Rich Domain 2 (STAC2) – chromosome 17q12

STAC2 is present in the nervous system and involved in skeletal muscle contraction.(Polster et al., 2015, Legha et al., 2010) No association with cancer has been made.

Hedgehog Interacting Protein (HHIP) – chromosome 4q31.21

HIPP regulates sonic hedgehog ligand signalling by binding to SHH on the cell surface. It is upregulated in response to SHH signalling and plays a role in sequestering the ligand in order to prevent autonomous cell and nearby neighbour cells (termed non-cell autonomous) from undergoing Hh signalling. This allows for a graded Hh signal which is essential for neural tube patterning seen in embryogenesis.(Kwong et al., 2014) Reduced HHIP expression is seen in combination with SHH over expressed cancer states such as pancreatic cancer, gastric cancer and medulloblastoma whereby the promoter region of *HHIP* undergoes methylation.(Martin et al., 2005, Song and Zuo, 2014, Shahi et al., 2011) High levels seen in the morphoeic subtype may reflect the attempt at tackling an aberrant Hh signalling pathway. HHIP is also significantly (P=0.02) overexpressed

(3.7 expression, log2FC 1.4) in nodular tumours and so reflects a common initiator or trunk change on the proviso that mBCC is an extension of nodBCC. Nevertheless, both subtypes rarely metastasis and this may in part be due to the anti-angiogenic effect of HHIP.(Olsen et al., 2004)

Endothelin 2 (EDN2) – chromosome 1p34.2

EDN2 belongs to a family of potent vasoconstrictors, but has also shown to have a role as a chemokine, inhibiting neutrophils in high concentrations and plays a role in the hypoxic environment seen in tumours.(Ling et al., 2013) EDN2 expression has been noted in nodular (and superficial) BCC before and the authors went on to suggest that it is a downstream target of Hh signalling.(Tanese et al., 2010) NodBCC also has significantly (P=0.002) overexpressed (expression 3.3, log2FC 2.5) EDN2, and this data supports the aforementioned study. EDN2 has been shown to be an inflammatory factor that promotes demyelination in the central nervous system.(Yuen et al., 2013) Increased expression has been noted in breast and renal cancer.(Grimshaw et al., 2002, Bot et al., 2012)

Calcitonin-Related Polypeptide Beta (CALCB) – chromosome 11p15.2

In contrast to EDN2, CALCB is a vasodilator.(Franco-Cereceda et al., 1987) However, it does play a role in the central nervous system as a possible neurotransmitter. This family of hormones has been associated with acute leukaemia and thyroid carcinoma.(Schifter et al., 1986, Pfluger et al., 1988)

Glypican 3 (GPC3) – chromosome Xq26.2

This is a cell surface proteoglycan which plays an important role in growth as highlighted by the loss of function mutation in *GPC3* causing Simpson–Golabi– Behmel overgrowth syndrome.(Pilia et al., 1996) Moreover, GPC3 binds SHH and inhibits Hh signalling. The GPC3-shh complex then undergoes endocytosis, mediated by low-density-lipoprotein receptor-related protein-1 (LRP1), which is non-significantly (p=0.07) overexpressed in morphoeic tumour (log2Fc 1.7, expression 6.1). This suggests that GPC3 has a tumour suppressor effect in cancer, and is decreased in breast and ovarian cancer.(Lin et al., 1999, Peters et al., 2003) However, it is over expressed in other cancers, including hepatocellular and lung, especially in the squamous subtype.(Lin et al., 2012, Du et al., 2011, Yu et al., 2015b) The mechanism of its role in tumorigenesis is not clear, but

thought to relate to binding to Wnt, resulting in its subsequent activation.(Lin et al., 2012) GPC3 is significantly (p=0.01) expressed (expression 4.2, log2Fc1.1) in nodBCC too.

BCL2-Related Ovarian Killer (BOK) – chromosome 2q37.3

As a member of the BCL-2 family, it has pro-apoptotic behaviour by mediating p53-apoptosis.(Bartholomeusz et al., 2006) For example, it has an important role in foetal and adult ovarian tissue, inducing apoptosis during follicle maturation.(Jaaskelainen et al., 2010) It has shown to be induced after DNA damage and be an essential mediator of p-53 dependent apoptosis in neuroblastoma and breast cancer.(Yakovlev et al., 2004) Over expression in morphoeic tissue probably represents a response to the excessive DNA damage seen in the tumour. Downregulation has been noted in colorectal cancer and confers an advantage to the tumour.(Zeilstra et al., 2011)

Melanoma Antigen Family A4 (MAGEA4) – chromosome Xq28

MAGE antigens are uniquely present on tumour and germ cells, but not on normal somatic adult cells. Its expression in cancer varies and its presence is associated with a poorer clinical outcome in melanoma, NSCLC and multiple myeloma.(Caballero and Chen, 2009) MAGE-A4 is not significantly (p=0.13) expressed in nodular subtype (expression 2.1, log2FC 1.3) and supports data that its expression is related to a more aggressive subtype. Despite this, the presence of a cancer unique antigen, if recognised by the body, can promote tumour cell death by the body's own immune system. This had led to the potential for immunotherapy and the development of vaccines against the antigen.(Gunda et al., 2013, Shirakura et al., 2012)

4.5.2 Downregulated genes in morphoeic BCC

Keratin 13, Type I (KRT13) – chromosome 17q21.2

Expression normally occurs in epithelial cells and in disease, it is associated with white plaques in the oral mucosa known as white sponge naevus syndrome. KRT13 was shown to be downregulated in oral/oesophageal squamous cell carcinoma and bladder cancer. (Worst et al., 2014, Hartanto et al., 2015) It is thought to be a marker of epithelial differentiation, whereby mild dysplastic tissue or well differentiated tumours retain KRT13 expression whereas, when tumours that become poorly differentiated they lose

expression. (Bloor et al., 2001) KRT13 is non-significantly (p=0.12) reduced (expression 3.1, log2FC -3.0) in nodBCC.

Keratin 4, Type II (KRT4) – chromosome 12q13.3

KRT13 and KRT4 are co-expressed in the same differentiated epithelial layers and unsurprisingly are both downregulated.(Sakamoto et al., 2011) Both are down regulated in squamous cell carcinoma of the larynx. (Nair et al., 2015) KRT4 is non-significantly (p=0.15) downregulated (expression 3.0, log2FC-2.1) in the nodular subtype. This biologically confirms the view seen on microscopy that morphoeic subtype is more poorly differentiated compared to the nodular subtype.

Transglutaminase 2 (TGM2) – chromosome12q11.23

Although originally described as having enzyme function in the extracellular matrix, it also plays a role in cell survival and signalling by binding to guanosine triphosphate (GTP). In breast cancer, TGM2 can induce epithelial to mesenchymal transition (EMT) and is involved in cancer stem cell traits.(Agnihotri et al., 2013) Overexpression confers a poor prognostic outlook and is associated with chemoresistance.(Leicht et al., 2014) This would support the low risk of metastasis seen in BCC. NodBCC demonstrates no change in TGM2 (P=0.89, expression 2.4 and log2Fc 0.2). In addition, TGM2 activates ROCK kinase, thus its deficiency promotes protein kinase B (AKT) signalling and adipogenesis.(Myneni et al., 2015) In addition, ROCK is involved in the axonal guidance pathway, a novel BCC pathway, and lack of ROCK has been shown to permit the survival of stem cells.(Watanabe et al., 2007)

Chemokine (C-X-C Motif) Ligand 17 (CXCL17) – chromosome 19q13.2

CXCL17 is a chemokine that attracts monocytes, dendritic cells and promotes angiogenesis.(Pisabarro et al., 2006) Monocytes are attracted by ERK1/2 as well as p38, and release proangiogenic factors including VEGF A. In contrast to other chemokines, it plays an anti-inflammatory role in LPS activated macrophages.(Lee et al., 2013) MBCC is often encased with a florid inflammatory reaction, and the loss of CXCL17 may augment this inflammatory response. Upregulation occurs in breast cancer and it confers a poorer prognosis in hepatocellular cancer.(Weinstein et al., 2006, Li et al., 2014a) CXCL17 has no significant change (p=0.91) in expression (expression 2.5, log2FC 0.1) in nodBCC, which lacks any surrounding inflammatory response.

Polymeric Immunoglobulin Receptor (PIGR) – chromosome1q32.1

PIGR is an important defence mechanism of mucosal surfaces that binds immunoglobulin (Ig) M and IgA, the latter being secreted as part of the innate immune system. Furthermore, the Ig/PIGR complex is secreted together to act as a free scavenger to attack microbes.(Phalipon and Corthesy, 2003) PIGR expression is reduced in colon cancer, an area where it is normally highly expressed as part of the colonic epithelial defence.(Traicoff et al., 2003) Reduced expression signifies tumour progression and poorer prognosis in pancreatic adenocarcinoma, more advanced tumour stage (but not metastasis) in gastric adenocarcinoma and less favourable prognosis in ovarian cancer.(Fristedt et al., 2014a, Fristedt et al., 2014b, Berntsson et al., 2014) Higher expression confers a protective effect as it inhibits the pro-inflammatory cytokine IL-8 which in turn activates polymorphonuclear neutrophils (PN) that are anti-tumorigenesis. (Dong et al., 2005) In addition, PN are known to activate matrix metalloproteinase 2 (MMP2) which aids local tumour invasion, a trait characteristic of morphoeic tumours.(Hadler-Olsen et al., 2013) NodBCC demonstrates non-significant (p=0.612) expression (log2FC 0.8, expression 3.3) of PIGR.

Immunoglobulin Lambda Variable 1-40 (IGLV1-40) – chromosome 22q11.22

Neoplastic plasma cells can produce an excess number of Ig light chains in conditions like myeloma. Polyneuropathy, Organomegaly, Endocrinopathy, Monoclonal gammopathy, and Skin changes syndrome (POEMs) is a rare disease of the plasma cell that preferentially releases light chain Ig. IGLV1-40 is also significantly (p=<0.001) downregulated (expression 2.9, log2FC -2.3) in nodBCC.

Immunoglobulin Kappa Variable 3-11 (IGKV3-11) – chromosome 2p11.22

IGKV3-11 is also significantly (P<0.001) downregulated (expression 1.7, log2FC -2.8) in nodBCC.

Membrane-Type Serine Protease 2 (TMPRSS4) – chromosome 11q23.3

This belongs to the family of serine proteases which are found throughout all cells and whose role is to cleave peptide bonds within proteins. Highly expressed TMPRSS4 induces invasion and migration in colon, pancreatic and gastric cancer.(Wallrapp et al., 2000, Kebebew et al., 2005) Moreover, high levels increases cell adhesion, migration and induce EMT via ZEB2 induction.(Jung et al., 2008a) Activation of invasiveness and

migration occurs via integrin alpha-5 (ITGA5) to trigger downstream pathways include FAK/MAPK, ERK, PI3K/Akt, Src and Rac1, with activation of the former two also being responsible for invasiveness and EMT.(Kim et al., 2010b) This has led to the development of TMPRSS4 inhibitors derived from 2-hydroxydiarylamide that has shown promise in colorectal cancer cells.(Kang et al., 2013) In contrast, lower levels (as seen in mBCC) were seen in metastatic melanoma: Riker et al compared melanoma profile to non-metastatic RNA profile of SCC and BCC. Although they did not state the subtype of BCC that were included in the study, they noted that all 15 had high expression of TMPRSS4.(Riker et al., 2008) However, TMPRSS4 was non-significantly (p=0.386) downregulated (expression 3.8, log2Fc -1.3) in nodBCC and so this also conflicts with Riker's findings.

Immunoglobulin Lambda Constant 2 (Kern-Oz- Marker) (IGLC2) – chromosome 22q11.22

Loss of heterozygosity (presumed reduce function) was noted in eight out of 11 intracranial meningioma and thought to play a pathogenic role.(Kim et al., 1993)

Immunoglobulin Heavy Constant Alpha 1 (IGHA1) – chromosome14q32.33

Differential expression has been noted in gastric cancer, but the authors did not mention if it was up or down-regulated.(Rajkumar et al., 2010)

Secretory Leukocyte Peptidase Inhibitor (SLPI) – chromosome 20q13.12

As the name suggest, SLPI protects epithelial cell surfaces from proteolytic enzymes (serine proteases) and is involved in the immune response, including being present in many secretions. Higher expression levels of SLPI induces cell proliferation and invasion in pancreatic ductal adenocarcinoma.(Zhang et al., 2015c) SLPI increases cell adhesion, Grb2, Ras and ERK1/2 phosphorylation.(Jeong et al., 2015) In addition, it is an anti-inflammatory molecule and inhibits the degradation of connective tissue by preventing the release of MMP2 and MMP9 by monocytes.(Zhang et al., 1997) Morphoeic BCC cells induce a significant inflammatory response and the loss of SLPI may contribute to this. MMP2 is non-significantly (p=0.43) expressed (expression 8.1, logFc0.6) in morphoeic and nodular (expression 8.1, log2FC -0.878, p=0.123) subtypes.

4.5.3 Upregulated genes in nodular BCC

Calpain 6 (CAPN6) – chromosome Xq23

Calpains are a large, ubiquitously expressed family of calcium dependent cysteine proteases, however, CAPN6 is mainly expressed in the placenta.(Dear et al., 1997) It is involved in microtubule stability and cytoskeletal organisation in embryo.(Tonami et al., 2007) Loss of CAPN6 promoted skeletal muscle differentiation, suggesting that it has an inhibitory role in skeletal muscle development.(Tonami et al., 2013) Increased expression is seen in uterine sarcoma and hepatocellular cancer.(Lee et al., 2007, Liu et al., 2015e) CAPN6 ability to inhibit apoptosis, via the PI3K-Akt pathway, may be the advantageous role it plays within the cancer.(Liu et al., 2011c)

Hsa-miR-3117– chromosome 1p31.3

Over-expression has been seen in colorectal cancer, more so in the primary tumour rather than metastatic tumour, and in melanoma.(Neerincx et al., 2015, Stark et al., 2010) This suggests it has a role in primary tumorigenesis and less so in the tumours ability to metastasis.

Zinc Finger Protein 418 (ZNF418) – chromosome 19q13.43

Krüppel or Cys2His2 zinc finger proteins are the largest class of transcription factors and ZNF418 acts as transcriptional repressor of the MAPK signalling pathway.(Li et al., 2008a) Overexpression of ZNF418, may therefore confer a disadvantage to the tumour.

Tolloid like-1(TLL1) – chromosome 4q32.3

During development, TLL1 is necessary for the correct position of the heart.(Clark et al., 1999) It is often co-expressed with BMP1 and when both are down-regulated, brittle bones (osteogenesis imperfecta) arise in the mouse model.(Muir et al., 2014) Expression of BMP1 enhances the migration of NSCLC cancer by releasing MMP2 and 9 via TGFB.(Wu et al., 2014) BMP1 is significantly (P=0.015) down regulated (log2Fc -2.5, expression 4.0) in nodBCC, but non-significantly (p=0.765) changed (log2FC 0.45, expression 4.0) in mBCC.

(RP11-119F7.5) – chromosome 10q22.1

This is a 2295 base pair, long non-coding RNA. Down-regulation of RP11-119F7.5 occurs in gastric cancer and is correlated with histological type and grade, but not survival prediction.(Sun et al., 2015)

(RP11-996F15.2) – chromosome 12p11.22

RP11-996F15.2 has 7 exons at 2380 base pair in length and has been associated with lymphoma.

Wiskott Aldrich Syndrome (WAS/WASL) Interacting Protein Family, Member 3 (WIPF3) – chromosome 7p14.3

WIPs play an important role in maintaining actin cytoskeleton integrity, cell adhesion, migration, chemotaxis and axonal growth.(Zettl and Way, 2002, Anton et al., 2002, Banon-Rodriguez et al., 2013, Lanzardo et al., 2007) WIP responds to B1 integrin to form complexes with CDC42 for fibroblast chemotaxis.(King et al., 2011) WIPF3 plays a role in spermatogenesis and the formation of a blood testis barrier.(Fu and Xiang, 2012) In cancer, it is found to have a role in breast cancer invasion probably via actin reorganisation.(Garcia et al., 2014)

Zinc Finger Protein 737 (ZNF737) – chromosome 19p12

ZNF737 resides in a cluster of genes on chromosome 19 whose expression is higher in human T-lymphocytes.(Bellefroid et al., 1993)

Thyroid Stimulating Hormone Receptor (TSHR) – chromosome 14q31.1

TSHR transmits the activity of the thyroid stimulating hormone (thyrotropin) onto the thyroid gland to release thyroxine and triiodothyronine. Gain of function mutations or hyper-expression is seen in thyroid cancer; with the mRNA levels acting as a biomarker for malignancy and can predict metastasis or residual disease.(Chia et al., 2007) Over expression has been shown to occur in the more aggressive form of breast cancer and may be associated with a poorer prognosis.(Yuan et al., 2014) *TSHR* mutation and over-expression is seen in some cases of lung cancer.(Kim et al., 2012)

Shisa family member 2 (SHISA2) – chromosome 13q12.3

SHISA2 regulates the Wnt and FGF pathways by modifying the number of frizzled and FGF receptors respectively.(Hedge and Mason, 2008) Both pathways play a role in gradient somitogenesis, which is promoted by SHISA2.(Nagano et al., 2006) It is up-regulated in breast, prostate and hepatocellular cancer cell lines and specifically related to more invasive behaviour.(Cheishvili et al., 2015)

4.5.4 Downregulated genes in nodular BCC

Solute Carrier Family 39 (Zinc Transporter), Member 2 (SLC39A2) – chromosome 14q11.2

SLC39A2 normal function involves the homeostasis of iron, zinc and calcium in developing embryos and is specifically expressed in keratinocytes, hepatocytes and dendritic cells.(Peters et al., 2007) SLC39A2 carries out contact inhibition of epithelial cells and so acts as a tumour suppressor. Damaging polymorphisms occur in this gene are related to bladder cancer.(Karagas et al., 2012) Low or undetectable levels have been noted in prostate and breast cancer.(Franklin et al., 2003) Zhang et al showed that breast cancer growth can be inhibited by increasing levels of Zn transporter LIV-1, but not SLC39A2, using flaxseed (the richest source of the plant lignin, secoisolariciresinol diglycoside).(Zhang et al., 2008)

Diacylglycerol Kinase, Theta 110kDa (DGKQ) – chromosome 4p16.3

The family of diacylglycerol kinases metabolise 1,2 diacylglycerol to produce phosphatidic acid. Expression of DGKQ is primarily seen in the nervous system where it plays a major role in synaptic transmission, gene expression, nerve growth factors and trafficking via various signalling pathways.(Tu-Sekine et al., 2013, Tabellini et al., 2004) Localised mainly in the nucleus, DGKQ attenuates protein kinase C (PKC) activity by removing diacylglycerol PKC activation and prevents progression of the cell cycle.(Sakane and Kanoh, 1997) Inhibitors to DCKs are being developed, but not specifically for the DGKQ isoforms, however these non-specific inhibitors have shown some promise.(Purow, 2015, Tu-Sekine et al., 2013)

Keratin 6A, Type II (KRT6A) – chromosome 12q13.3

The keratin family are fibrous structural proteins responsible for cytoskeleton integrity. Keratin filaments are present in epithelial cells and especially prominent on the surface

of skin and found in nails, hair and hard palate of the mouth. Autosomal dominant mutations in the *KRT6A* gene result in pachyonychia congenita type1 (MIM 167200), a condition characterised by hypertrophic nail dystrophy, palmoplantar keratoderma, follicular keratosis and oral leucokeratosis.(Bai et al., 2008) KRT6A forms heterodimers with KRT16 and 17. The former is down-regulated (-2.9 log2FC, expression 6.5) significantly (p=0.004) in nodBCC and in mBCC (-1.6 log2FC, expression 6.5 p=0.005). KRT17, however, is non-significantly (p=0.28, 0.9 log2FC, 2.4 expression) changed in nodBCC whereas it is significantly (p=0.03) upregulated (log2Fc 1.3, expression 2.4) in mBCC. KRT6A negatively regulates the proto-oncogene tyrosine-protein kinase (SRC) and loss of expression enhances keratinocyte migration.(Rotty and Coulombe, 2012)

Cytohesin 4 (CYTH4) – chromosome 22q13.1

Cytohesins are a small family of four and a type of guanine nucleotide-exchange protein (GEP).(Moss and Vaughan, 2002) They mediate phosphoinositide (PI) 3-kinase regulated ADP ribosylation factor (Arf) signalling.(Jackson et al., 2000) CYTH4 expression is localised to leucocytes, especially monocytes (CD33+) and absent from most other tissue.(Ogasawara et al., 2000) The lack of CYTH4 may reflect the tumours ability to evade the monocyte innate immune system and their progeny of macrophages and dendritic cells.

CD2 Molecule (CD2) – chromosome1p13.1

All peripheral blood T-lymphocytes possess a surface antigen CD2 along with some natural killer cells. In cutaneous melanoma, CD2 cell number has a prognostic role, with fewer cells being a higher risk of recurrence or death.(Harcharik et al., 2014) A similar result was seen in T-cell acute lymphoblastic leukaemia whereby CD2-negative patients had a poorer outcome.(Uckun et al., 1996) MBCC is more aggressive than nodBCC, but the down-regulation of CD2 is less (log2FC -1.3, expression 4.5), but still significant (P=0.05). Nevertheless, it does highlight the fact the adaptive immune system is also dysregulated in BCC.

Histidine Decarboxylase (HDC) – chromosome 15q21.2

This is the sole enzyme that provides histamine in the body by converting L-histidine into histamine, hence is a biomarker for histamine synthesis. It is a well-known inflammatory mediator, but also involved in neurotransmission, secretions and smooth muscle tone.

Increased expression is seen in colorectal, pancreatic, small cell lung cancer and melanoma.(Masini et al., 2005, Tanimoto et al., 2004, Matsuki et al., 2003, Haak-Frendscho et al., 2000) Histamine released from mast and non-mast cells induce tumour angiogenesis.(Li et al., 2008b) In addition, it seems to augment an autocrine tumorigenic loop in cholangiocarcinoma by enhancing local histamine synthesis.(Francis et al., 2012). It is non-significantly (P= 0.19) down-regulated (-1.7 log2FC, expression 2.6) in mBCC too and therefore histaminergic tumorigenesis does not seem to play a part in BCC.

Family with Sequence Similarity 65, Member C (FAM65C) – chromosome 20q13.3

Over-expression has been seen in cervical, pancreatic, renal, large bowel and urinary tract cancer, with the latter in over 10% of cases.(Forbes et al., 2011)

Guanylate Binding Protein Family, Member 6 (GBP6) – chromosome 1p22.2

Common interferon stimulated genes include the guanylate binding proteins as part of the innate immune system. Loss of GBPs result in defective macrophage activity.(Pilla et al., 2014) GBP1 expression is associated with a better prognosis and inhibits proliferation, invasion and migration of colorectal cancer cells, therefore, GBP6 may act as a tumour suppressor gene.(Britzen-Laurent et al., 2013)

Granulysin (GNLY) – chromosome 2p11.2

T-lymphocytes and natural killer cells contain cytotoxic granules called granulysin as part of the innate immune system which is lytic to tumour cells.(Stenger et al., 1999) Increased expression conferred a favourable outcome in gastric cancer, colorectal cancer and large B-cell lymphoma.(Li et al., 2012b, Tosolini et al., 2011, Park et al., 2011) It induces apoptotic cell death in haematological tumours and its anti-tumour properties could be exploited as a form of immunotherapy.(Aporta et al., 2014)

CD1c Molecule (CD1C) – chromosome 1q23.1

Four transmembrane glycoproteins for the family of CD1 molecules that are involved in the major histocompatibility complex and the presentation of lipid antigens to T-lymphocytes. CD1C is known to be involved in antigen presentation in relation to *Mycobacterium tuberculosis*.(Matsunaga and Sugita, 2012) It traffics through all cell compartments unlike the other isoforms and presents a wide range of antigens to T-cells including endogenous antigens, hence may play a role in cancer cell priming of the body's

T-cell defence.(Adams, 2013) CD1C is a marker of professional antigen presenting cells such as the dendritic cell and encouraging its ability to present cancer cell lipids as another form of immunotherapy.

4.5.5 Upregulated in morphoeic compared to nodular BCC

Plakophilin 1 (PKP1) chromosome 1q32.1

Being an armadillo-related protein, it has a dual role in cell-to-cell contact and cytoplasmic/nuclear signalling altering gene expression. PKP1 acts both in the extracellular matrix (PKP1a) to support desmosome structure and as a transcription factor in the nucleus (PKP1b variant) to promote eukaryotic initiation factor 4A1 (eIF4A1) activity.(Wolf et al., 2010) Deficiency in PKP1 was first noted in the dysplasia-skin fragility syndrome whereby an autosomal recessive mutation leads to skin fragility, palmoplantar hyperkeratosis, onychodystrophy, perioral fissuring and non-cicatricial alopecia due to lack of protein production.(Hernandez-Martin et al., 2013) A shortage of PKP1 results in a reduction in the number of desmosomes and increased migration.(South et al., 2003) Plakophilin expression in breast cancer demonstrates overexpressed PKP3, but not PKP1 or 2.(Demirag et al., 2012) Gastric cancer showed a loss of PKP3, but no change in PKP1 or 2.(Demirag et al., 2011) Oesophageal adenocarcinoma shows methylated *PKP1* (therefore reduce PKP1) resulting in increased cell motility.(Kaz et al., 2012) Cervical cancer demonstrates a reduced expression of PKP1.(Schmitt-Graeff et al., 2007) Conversely, increased expression is seen in squamous cell carcinoma, especially in those at low risk of metastasis. (Villaret et al., 2000) Differing plakophilin expression may be due to them having a dual role as either an oncogene or tumour suppressor gene depending on the cancer cell of origin.(Bass-Zubek et al., 2009) Moreover, cytoplasmic PKP1 may act as a post-transcriptional regulator of gene expression. (Fischer-Keso et al., 2014) PKP1 immunostaining of BCC and SCC have been heterogeneous, with a trend towards less intensity in poorly differentiated (and invasive) tumours.(Moll et al., 1997) In contrast, PKP1 in nodBCC is significantly ($p=0.002$) downregulated (expression 7.3, log2Fc -1.6).

Laminin, Beta 3 (LAMB3) – chromosome 1q32.2

The laminins are a family of extracellular proteins involved in a diverse number of functions depending on the organ. In skin they are involved in cohesion of the dermal-

epidermal junction and mutations in *LAMB3* (homozygous loss of function) result in blistering skin disorders such as epidermolysis bullosa.(Laimer et al., 2010) Laminins have been shown to be involved in cancer and may aid local spread or invasion. (Sroka et al., 2010)

Glypican 1 (GPC1) – chromosome 2q37.3

There are six glypicans within the family and two main subgroups based on sequence homology. GPC1, 2, 4, 6 are in one subgroup and 3, 5 in the other. As mentioned, they have an important role in developmental morphogenesis and interact with Hh, Wnt, FGF and BMP.(Filmus et al., 2008) GPC 1 regulates Hh signalling by acting as a co-receptor in Shh-dependent induction of commissural axon guidance.(Wilson and Stoeckli, 2013) Its responsibility in cancer is less well documented than its family member GPC3, however, it is overexpressed in pancreatic cancer, and is part of the stromal soup of MMP, growth factors that promotes invasion of cancer cells.(Korc, 2007) It is also overexpressed in ameloblastoma, glioma and breast cancer.(Bologna-Molina et al., 2015, Qiao et al., 2013, Matsuda et al., 2001) GPC1 inactivates the G1/S checkpoint and strongly stimulates DNA replication in glioma cells, possibly via the S-phase kinase-associated protein 2 (SKP2) autoinduction loop.(Qiao et al., 2013) In pancreatic cells, GPC1 is essential for the mitogenic effect of FGF2 and HB-EGF.(Kleeff et al., 1998)

EPH (Ephrin) Receptor B4 (EPHB4) – chromosome 7q22.1

Ephrins mediate developmental processes, especially in the nervous system. Both EPHB3 and EPHB4 are significantly upregulated (the latter log2FC 2.0, expression 3.6 and $P<0.003$). EPHB4 is a receptor tyrosine kinase that coordinates cell migration during embryogenesis and although ubiquitous in adult tissue, is expressed at low levels. Overexpression has been detected in lung, colorectal, breast, thyroid, prostate, cervix, ovarian and skin cancers. EPHB4 promotes proliferation, cell motility and migration which is enhance by binding of the ligand ephrinB2.(Ferguson et al., 2014) The receptor binds all ephrin-B ligans and EFNB1 is significantly ($P=0.007$) overexpressed (1.2 log2FC, expression 4.7) in mBCC, but non-significantly ($p=0.559$) altered (0.3 log2FC, expression 4.7) in nodBCC. Mechanism of action include targeting Rho GTPases, directly promoting angiogenesis by action on endothelial cells and affecting the extracellular matrix by modulating integrins such as Integrin beta 8 (ITGB8) to aid cell

migration.(Yang et al., 2006, Mertens-Walker et al., 2015) Inhibitors of EPHB4 are in the pipeline and clinical trials have already started with the use of EPHB4-fusion protein and small molecules in combination with standard chemotherapy for solid tumours.(Duggineni et al., 2013)

Unique to morphoeic tumour and normal stroma

The above-mentioned genes are DE when comparing tumour subtypes only. When looking at all four compartments (two tumour subtypes and two normal stroma) we can look for genes that are shared (see 4.2.4iii below) and those that are present in some of the compartments. DDR1 was found to be expressed in both the morphoeic tumour and normal stroma, but not the nodBCC tumour or the corresponding normal stroma. For example, comparing morphoeic tumour with nodular normal stroma, the expression was 7.2, log2FC 1.8 and P 0.0005. This raises the possibility of the normal surrounding stoma in the morphoeic tissue being different to true, normal tissue and DDR1 is discussed further below. In order to look at this further, the morphoeic normal stroma was compared to the nodular normal stroma to identify any other interesting genes.

Discoidin Domain Receptor Tyrosine Kinase 1 (DDR1) - chromosome

DDR1 is a tyrosine kinase receptor that binds collagen to induce cell migration, proliferation and extracellular matrix modification. Pathway activation is dependent on cell type as it can activate or inhibit ERK pathway.(Curat and Vogel, 2002, Lu et al., 2011) Within the extracellular matrix it can activate integrins, TGFB and notch1.(Kim et al., 2011) Overexpression has been shown in pancreatic, NSCLC, breast, ovarian and hepatocellular cancer.(Huo et al., 2015, Shen et al., 2010b) DDR1 Has been shown to mediate MMP invasion in breast cancer cells and regulates embryonic stem cell self-renewal in the mouse.(Castro-Sanchez et al., 2011, Suh and Han, 2011) Inhibitors of DDR1 have been developed that reduce tumorigenesis and repel metastatic cells from invading bone.(Valencia et al., 2012, Elkamhawy et al., 2015, Kim et al., 2013b)

4.5.6 Shared genes present in both morphoeic and nodular tumour

SH3 And Cysteine Rich Domain 2 (*STAC2*) chromosome 17q12

STAC2 is part of a membrane trafficking adapter protein family and one of three isoforms (STAC 1-3). It is found in skeletal muscle, spine, cerebellum and forebrain.(Nelson et al., 2013) In skeletal muscle, it increases calcium influx and subsequent contraction.(Polster et al., 2015) STAC2 is also expressed during post-natal development of dorsal root ganglion neurons and promotes their expansion.(Legha et al., 2010) Potentiating growth may be the role that it plays in cancer.

Versican (*VCAN*) – chromosome 5q14.2

The extracellular matrix within the developing embryo provides a permissive environment for growth and VCAN is a proteoglycan that forms a part of this complex. In addition, it is found in large amounts in adult tissue with particular prominence in soft tissue.(Wight, 2002) Protein Tyrosine Kinase 2 (PTK2), PDGF and VCAN all work together to expand the extracellular matrix for proliferation of cells.(Wight, 2002) PTK2 is non-significantly (p=0.27/0.19) downregulated (-0.5/0.6 log2FC, expression 5.8/5.8) in morphoeic/nodular respectively, but PDGFA is overexpressed (see below). It is proinflammatory by allowing sustained recruitment of leukocytes and release of inflammatory chemokines and helps to create a setting for cell proliferation and migration. VCAN is upregulated by PDGF and TGFB1, both known mitogens, and it is possible that VCAN itself is a mitogen by acting through EGF.(Wight, 2002) It terms of cell migration, it has been extensively studied in relation to the movement of neural crest cells during embryogenesis.(Perissinotto et al., 2000) Increased expression has been seen in a variety of cancers including breast, prostate, ovarian and confers an inverse relationship with prognosis.(Ween et al., 2011, Du et al., 2013) Specifically, it plays a role in the local stromal microenvironment in association with hyaluronan (HA) and CD44.(Ricciardelli et al., 1997) Immunostaining for VCAN in BCC demonstrated stromal expression, but not within the tumour; HA and CD44 showed very low tumour cell signal.(Karvinen et al., 2003) Promising results have been shown by the inhibition of VCAN with the tyrosine kinase drug Genistein, asthma drugs Montelukast and Formoterol and MMP inhibitors.(Ween et al., 2011)

Platelet-Derived Growth Factor Alpha Polypeptide (PDGFA) – chromosome 7p22.3

Platelet-derived growth factors are potent mitogens of mesenchymal cells and essential in embryogenesis, cell survival, chemotaxis, migration and proliferation. Overexpression is seen in cholangiocarcinoma, glioma, mantle cell lymphoma, pancreatic and ovarian cancer. (Martinho et al., 2009, Boonjaraspinyo et al., 2012) It has been demonstrated by immunohistochemistry that stromal PDGFA expression is necessary for BCC growth.(Ponten et al., 1994) Activation of PDGFA by SOX11 has been shown to regulate angiogenesis in mantle cell tumours with SOX11.(Palomero et al., 2014) Furthermore, SOX11 itself promotes angiogenesis in mantle cell tumours and effective tumour growth by activating AKT and MAPK. Sox11 is significantly ($P=0.001/P<0.001$ overexpressed (1.8/2.3 log2FC, expression 5.3/2.3) in morphoeic/nodular respectively. Clinical trials have been undertaken with anti-PDGFA, however, it has not shown to be potent as a single agent, rather, used in combination with chemotherapy.(Raymond et al., 2008)

Basonuclin 2 (BNC2) – chromosome 9p22.2

BNC2 is virtually ubiquitous to the nucleus of all cell types and highly expressed in reproductive tissue.(Winham et al., 2014) It has been shown to regulate male mouse stem cells, allowing for meiosis progression in spermatogenesis.(Vanhoutteghem et al., 2014) In oesophageal cancer it has been suggested to be a tumour suppressor gene, based on expression of BNC2 causes growth arrest in vitro.(Akagi et al., 2009) Basonuclin 1 has been shown to be overexpressed in BCC and upregulated by Hh/Gli (especially Gli2) signalling, however, despite the similarity of name, the protein function is vastly different to BNC2.(Vanhoutteghem and Djian, 2006, Cui et al., 2004) Moreover, BNC1 is not over or under expressed in either nodBCC or mBCC.

Glypican 3 (GPC3) – chromosome Xq26.2

Please refer to section 4.2.2i

Adenomatous Polyposis Coli membrane recruitment protein 2 (AMER2) – chromosome 13q12.13

Please refer to section 4.5.1

Shared progressive gene

CTSV was found to be present in all four compartments (i.e. both the normal stroma and tumour of the two subtypes), but had increased in its expression in a progressive fashion starting with the lowest expression in normal nodular tissue, then increasing in the nodBCC tumour, morphoeic normal tissue and highest within the mBCC tumour. It is discussed in further detail below.

Cathepsin V (CTSV) – chromosome 9q22.3

Cathepsins form a large family proteases, and CTSV is a cysteine proteinase involved in collagen degradation, keratinocyte differentiation and angiogenesis.(Zeeuwen et al., 2007) The pathogenesis of fibrosis, vasculopathy and altered keratinocyte function seen in systemic sclerosis has been attributed to loss of CTSV expression.(Noda et al., 2013) Expression of CTSV is under the control of Fli1, and this is just non significantly (P=0.06) downregulated (log2FC-1.11, expression 4.3) in nodBCC, but significantly (p=0.02) downregulated (log2FC -1.4, expression 4.3) in mBCC.(Noda et al., 2013) Cathepsin K, also a cysteine proteinase, was expressed in BCC and Bowen's disease, and cathepsins in general are associated with cancer. (Ishida et al., 2013, Loser and Pietzsch, 2015)

CONCLUSION

4.6 Summary of gene expression in BCC

Basal cell carcinoma (BCC) is the most common cancer in the world and a large proportion (60%) affect the face.(Flohil et al., 2013) In particular, BCCs within the H zone, including periocular region, have a higher risk of recurrence, behave more aggressively and pose a risk to sight (see Figure 1.6).(Mosterd et al., 2008) Compounding this is a histological diagnosis of morphoeic BCC (mBCC), which is a high risk factor for recurrence and often requires Mohs micrographic surgery to ensure complete histological resection.(Malhotra et al., 2004) mBCC comprises of irregular, spiky strands of cells accompanied by a florid surrounding inflammatory fibrosis of the stroma (Figure 1.7A) In contrast, the more common, indolent nodular BCC (nodBCC) consists of regular nests of cells, a pseudocapsule and no surrounding stromal inflammation (Figure 1.7B).

Whole exome sequencing (WES) of 10 periocular mBCC were compared to 10 nodBCC; 5 periocular and 5 below the head to allow for histological and site specific evaluation. BCCs in the periocular region were hypothesised to have a greater mutational burden and higher UV burden than BCC elsewhere (elseBCC) with mBCC being the most genetically altered. Overall mutational burden was marginally (p=0.08) less in mBCC than nodBCC (Table 3.2), but locality had no effect (P=0.6) with normal skin burden be a confounding factor. (Martincorena et al., 2015)

Driver mutations in mBCC using the Intogen algorithm ($q < 0.1$) revealed 18 significantly mutated genes, including two known drivers *PTCH1* and *SMARCA4*, four Intogen cancer drivers *CHD3*, *EFTUD2* and *ARHGAP35,* and additional genes shown to be connected to a driver, such as *FLNB, EPHA3* and *CHKB* (Table 3.4). MutSigCV identified only one known driver *PTCH1* ($q < 0.05$), with additional top mutated genes being *CCDC108*, *FBX015*, *POM121L12* and *OR5H2* (raw $p < 0.02$) (Table 3.14). Shared genes across the two algorithms revealed *PTCH1* only (Figure 3.9).

In contrast, in nodBCC exomes Intogen highlighted 23 significantly mutated genes ($q < 0.05$), including four known drivers *PTCH1, TP53*, *MYCN, ARID1A* and five Intogen drivers *ARHGAP35*, *IRF2*, and *PPM1D* (Table 3.15). MutSigCV revealed 11 top mutated

genes ($p < 0.02$) with the top two being *TP53* and *PTCH1* (Table 3.14). Shared genes across the two algorithms uncovered six drivers, *PTCH1*, *TP53*, *FADS1, PPM1D, TAF4B and TRIP4* (Figure 3.13).

Next, we performed the randomisation analysis (see Methods) to identify significantly mutated genes that were related to a specific subtype or locality. Intogen mBCC specific genes included *HECTD4* and *FLNB* and nodBCC specific were *ZEB1* and *TP53* (Figure 3.17). MutSigCV mBCC specific gene included *CCDC108*, and nodBCC included *GABRA6*, *OR52J3*, *TRIM39* and *ADAM29* (Figure 3.14). Locality nodBCC genes include *SIPA1L1* and *SYNE1* whereas elseBCC revealed *ANK2* and *HIVEP1* (Figure 3.18). *TP53* and *PTCH1* were shared across both locations for nodular BCC.

Bonilla *et al* recent article revealed seven driver mutations in their large collection of BCCs containing mixed histological subtypes and we demonstrated a similar burden in our 20 cases *PTCH1* (75%), *TP53* (45%), *SMO* (15%), *MYCN* (25%), *PTPN14* (10%), *RPL22* (20%) and *PPIAL4G* (10%) although these mutations may reflect a nodBCC makeup which comprises the bulk of Bonilla's cases.

Tumour-stroma paired gene-set enrichment analysis (GSEA) on RNAseq expression was performed and compared to Intogen pathway prediction on WES data. Four common mBCC dysregulated pathways ($q < 0.05$ for both WES and RNAseq) were; 'hedgehog (Hh) signalling pathway','BCC','Natural killer cell mediated cytotoxicity' and 'Fc Epsilon RI signalling pathway' (Figure 4.11). In nodBCC, the 'BCC', 'PPAR signalling' and 'oxidative phosphorylation' were commonly dysregulated (Figure 4.13). Interestingly, Hh, TGF-beta and p53 signalling demonstrated significant mutational burden that was not reflected in the differential expression (DE) profile in nodBCC. The comparison between mBCC and nodBCC WES only data revealed a significant novel shared pathway as 'Wnt signalling', in addition to 'BCC', 'Hh signalling', 'pathways in cancer' and 'cell cycle'. Comparison of nodBCC and mBCC GSEA only revealed multiple shared immune-related pathways including 'Natural killer cell mediated cytotoxicity' (Figure 4.8). However, significant upregulation in 'Hh signalling', 'spliceosome' and 'base excision repair', and downregulation in 'ABC transporters' and 'glycosphingolipid biosynthesis' were observed to mBCC only. Pathway analysis by

Bonilla *et al* again reflected more nodBCC specific features with the presence of p53 and TGF-B signalling, something revealed by mutational burden only in our data.

Epithelial cancer behaviour is not restricted to the tumour borders delineated on microscopy; the surrounding stroma has been shown to play a key role in local spread(Davidson et al., 2014) (Figure 1.7). WES of one mBCC microdissected tumour-stroma pair revealed 150 shared nonsynonymous mutations (76.5% of all stromal mutations) including mutations in *ATR*, *EPHA3* and *Gli3* (Figure 3.20), although the overall variant allele frequency (VAF) was low in both samples, reflecting the low tumour purity. *EPHA3* is thought to act as a tumour suppressor gene and mutations have been found within the surrounding stroma.(Boyd et al., 2014, Vail et al., 2014) The highest mBCC tumour VAF genes include *ROS1*, a proto-oncogene tyrosine kinase receptor and Gli3, a Hh protein involved in cancer progression.(Wen et al., 2014) This raises the question of whether this is true stromal mutational burden or just histologically undetectable cancer cells intermixed within the stromal soup. Either way it is clinically relevant when it comes to surgical excision whereby the removal of this inflammatory material along with tumour should be considered, as failure to do so may be the reason for the local recurrence seen in mBCC.

Protein overexpression of Hh was seen in mBCC tumour compared to nodBCC tumour supporting the aforementioned RNAseq data (Figure 4.14). Stromal Hh protein expression was significantly up regulated too in mBCC whereas nodBCC was similar or less than control tissue (Figure 4.15). Activation of stromal Hh pathway could be a local paracrine effect from the morphoeic tumour or related to the altered cells in the stroma itself. Moreover, thirty-seven DE genes (raw $p < 0.01$) were seen in mBCC stroma compared to normal eyelid. *GSTM1*, a detoxification enzyme associated with BCC, was most significantly, upregulated in mBCC stroma with almost absent expression in nodBCC stroma. GSEA analysis revealed the most upregulated pathways in mBCC stroma were immune related, such as ‘primary immunodeficiency’ and ‘T-cell receptor signalling’. Analysis of Hh signalling expression in RNAseq also demonstrated a trend towards increased expression in mBCC stroma, although it was not significant compared to nodBCC stroma.

The driver mutational profile is different in mBCC compared to its indolent nodular counterpart. Loss of function *PTCH1* is common to both subtypes; however, their

markedly different behaviour may also be reflected by differing expression of the Hh pathway especially within the surrounding stroma tissue. Removal of both mBCC tumour and stroma may be required to prevent recurrence.

4.7 Potential treatment targets

The overarching aim to all of this research is to identify potential treatment targets for patients with BCC and after deciphering the data there are targets where there is a currently available treatment in humans:

Smoothened (*SMO*) inhibitor

Development of SMO inhibitors occurred after observations were made with the *Veratrum californicum* plant extract cyclopamine. Cyclopamine itself has a poor pharmacokinetic profile and too toxic for humans. Great promise was shown during one of the first trials in Gorlin syndrome whereby a 58% response rate was seen in advanced BCC on 150mg PO Vismodegib daily, nevertheless 25% reported serious side effects including 7 deaths.(Sekulic et al., 2012) Safety outcome reporting at 12 months since the trial demonstrated no more death adverse events but similar side effects of muscle spasm (72%), alopecia (66%) dysgeusia (54%), weight loss (50%), nausea (33%) and lethargy (41%).(Sekulic et al., 2015) Periocular mBCC tumours have been treated with Vismodegib, and a small case series of 7 advanced mBCC (post Mohs clearance relapse): Two achieved complete clearance, one continued to progress with the others partially responding, however, 86% experienced side effects as mentioned above, and the development of cutaneous squamous cell carcinoma in two cases.(Gill et al., 2013) Resistance to the Vismodegib has also been shown and this is thought to involve upregulation of PI3K. To combat resistance, combination therapy is currently being assessed with PI3K inhibitors in advanced/metastatic BCC (NCT02303041; Buparlisib + Sonidegib) and advanced solid tumours (NCT01576666).

Glioma-associated oncogene (Gli) zinc finger transcription factors 1/2 (*Gli1/2*)

Hh is an important signalling pathway in many cancers and as confirmed in our study, plays a pivotal. Although initial responses with SMO inhibitors were promising, resistance has now been observed, recurrence when stopping the treatment. Resistance was shown in advance BCC and medulloblastoma.(Yauch et al., 2009) Re-growth was

shown in 20% of patients whilst on SMO treatment.(Chang and Oro, 2012) The antifungal Iitraconazole (antifungal) was found to inhibit Hh downstream of SMO and thought to prevent Gli transcription.(Kim et al., 2010a) Furthermore, a clinical trial is recruiting at the national cancer institute, USA comparing the effects of Iitraconazole versus placebo (NCT02735356). Gli antagonists (GANT) have shown promise across a spectrum of animal cancer models including pancreas, lung, prostate and rhabdomyosarcoma. GANT are thought to act in a multitude of ways that are tumour suppressive (Fig 3.34). In acute myeloid leukaemia (AML) cells it causes growth arrest and apoptosis, further, in combination with a mTOR inhibitor (Rapamycin) it has shown synergistic potential.(Pan et al., 2012) Similar actions were detected in treating rhabdomyosarcoma, along with reduction in EMT transition vai AKT/mTOR signalling.(Srivastava et al., 2014) On a head to head with the SMO inhibitor Vismodegib, GANT61 was found to suppress tumour growth more effectively.(Benvenuto et al., 2016) This highlights the importance of targeting the downstream part of the signalling pathway, especially if there is suspected non canonical activation of Gli.

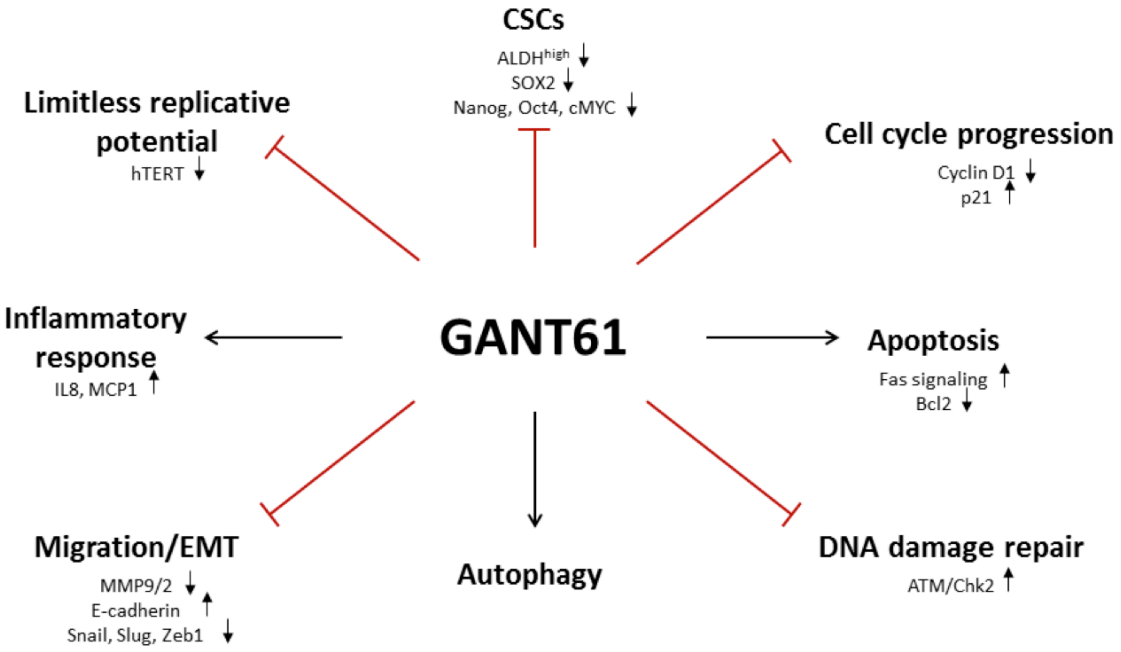


Fig 4.18 Action of Gli antagonists. Summary of proposed anti-cancer actions by Gli antagonists (GANT). Adapted from Pan *et al* 2012.

Versican (*VCAN*) inhibitor

Induction of VCAN has shown to benefit cancer progression, migration, invasion and associated with a poorer outcome. Genistein is a phenolic compound present in soy based foods and shown to have a wide variety of anti-cancer effects.(Spagnuolo et al., 2015) Nevertheless, it has been shown to increase proliferation of breast cancer cells at high doses, so there may be a narrow therapeutic window outside of which may be dangerous.(Russo et al., 2016) Leukotriene inhibitors reduce versican and are shown to reduce the risk of cancer development in asthma sufferers.(Tsai et al., 2015) In colonic cancer cells $CysLT_1R$ antagonists (for example Montelukast and ZM198,615) have been shown to reduce proliferation and colony formation.(Savari et al., 2013)

Wnt signalling targets

Protein-serine O-palmitoleoyltransferase porcupine (PORCN) interacts with wnt1,3,3A,4,5A,5B,6,7A,7B to carry out palmitoylation, a step required for ligand secretion. Small molecule, Wnt974 inhibitor (Novartis), has been developed that binds to PORCN and inhibits the enzyme. Initial findings in Wnt driven rodent tumour models, and in loss of function NOTCH signalling where cross talk occurs with Wnt, showed promising results.(Liu et al., 2013) The molecule has moved into clinical trial for pancreatic adenocarcinoma and BRAF mutated colorectal cancer. It is possible other tumours with variants in upstream Wnt signalling could be responsive. Monoclonal antibodies have been developed against the FZD receptor including Vanituctumab which binds to five of the ten frizzled receptors (FZD 1,2,5,7,8). It has shown promise in pancreatic and breast cancer cells, but works better in combination with chemotherapy and is currently undergoing a phase 1 clinical trial in the aforementioned cancers in humans. Cancer like stem cells (CSC) self-renewal depends of Hh ,Wnt, and notched signalling and they possess specific markers CD133, CXCR1, and CD44. CSC actions are attributed to relapse and metastasis via the c-met/FZD8 axis. An antibody specific to FZD8, omp-54 F28, has been used in humans but causes significant side effects in over 20% of patients including bone fractures. The latter complication would prevent it being a practical drug in metastatic bone disease where fractures are common, however it has progressed into phase 1b study as an adjunct in ovarian, hepatocellular and pancreatic cancer. Antagonists of the cAMP response element binding protein (CREB)-binding protein (CBP)/catenin prevent the interaction of the two proteins and in turn eliminate cancer initiating cells. By preventing their combination, it has been shown to reduce the

formation of drug resistance. These have gone on to phase 2 clinical trials in combination with avastin for metastatic colon cancer. Aberration of Wnt within cancer occurs at different levels of the pathway, for example, modification of FZD receptor function would not be helpful in APC colorectal cancer where the aberration arises lower down in the pathway. Traf2- and Nck-interacting protein kinase (TNIK) is an essential regulator of the β-catenin and T-cell factor 4 (TCF4) complex, the final component of Wnt signalling. Aminothiazole is a small molecule inhibitor of TNIK and shown to reduce Wnt signalling as have some other molecules in the preclinical setting.(Masuda et al., 2015) It does have a role in AKT pathway, autophagy and EMT so the effect of blocking these is unknown.

Histone deacetylase 1 (*HDAC1*) inhibitor

This inhibitor has been used in the past as a mood stabiliser and antiepileptic. Vorinostat was one of the first inhibitors on the market and licensed in 2006 against cutaneous T cell lymphoma. Panobinostat is one of the latest licensed HDAC inhibitor in 2015 and had been used in patients with haematological and solid cancers. Route of administration has been oral and intravenous although there is a wide variability between patients in active drug getting to the tissue of interest despite administered the same dose.(Srinivas, 2016)

Erythropoietin-producing hepatocellular carcinoma (EPH) Receptor A3 (*EPHA3*)–inhibitor

EPHA3 has large body of evidence that it is associated with cancer, even more that it is a driver mutation. It has a role in both the Wnt and axonal guidance pathway that are altered in mBCC. In nodBCC, only Wnt is altered. Its role is cellular specific and is deleted in some cancers and overexpressed in others. A monoclonal antibody IIIA4 triggers EPHA3 activation, preferentially enters tumour cells and a humanised version, KB004 (Kalabios Pharmaceuticals Inc, USA), causes apoptosis of EPHA3 leukaemia stem cells from acute myeloid leukaemia patients. In a phase 1 trial it reduced blast counts in blood and bone marrow cells.

EPH (Ephrin) Receptor B4 (EPHB4) inhibitor

As mentioned, antibodies have been developed against EPHB4 and promising results have been shown using monomeric soluble EphB4 fused to human serum albumin (EphB4-HSA) against human mesothelioma cells established with human mesothelioma

cells.(Savari et al., 2013) Phase 1 trial has been carried out in ovarian cancer using JI-101, a multi-kinase inhibitor against EPHB4, PDGFRB and VEGFR-2 in ovarian cancer.(Werner et al., 2015)

Anaplastic Lymphoma Receptor Tyrosine Kinase (ALK) inhibitor

Crizotinib (PF-02341066) is an oral inhibitor of ALK along with MET, ROS1 and RON kinase. It has been trialled in MET and ALK activated tumours

Hydroxysteroid (11-Beta) Dehydrogenase 1 (HSD11B1) inhibitor

Local tissue levels of active cortisol are controlled by 1β-hydroxysteroid dehydrogenase type 1 and 2 with production of cortisol, forward reaction by the former. Various compounds are under development primarily for the treatment of type 2 diabetes where it is thought that high levels of cortisol in adipose tissue cause visceral obesity and insulin resistance.(Hong et al., 2015)

Peptidyl Arginine deiminase type III (PADI3) inhibitor

Peptidylarginine deiminases are involved in the deamination of the amino acid arginine into citrulline within a protein, a process also known as citrullination. This process is believed to take part in various chronic inflammatory diseases such as rheumatoid arthritis, psoriasis fibrosis and Alzheimer's disease.(Gudmann et al., 2015) Proteins that undergo citrullination undergo degradation more quickly and in terms of tumour, may make the microenvironment more conducive to local spread. Analogues of benzoylarginine amide are the best PAD inhibitors.(Knuckley et al., 2010, Jones et al., 2009) Cl-amidine and F-amidine have been shown to reduce cancer cell viability and sarcoma growth in murine models.(Slack et al., 2011, Wang et al., 2012b)

ROS proto-oncogene 1 (*ROS1*) – chromosome 6q22

Treatment of ALK positive NSLC with specific antibodies has been developed including Crizotinib (Xalkori®, Pfizer) and some have activity against ROS1..

Melanoma Antigen Family A4 (MAGEA4) – chromosome Xq28

The development of vaccines against this antigen has occurred .(Gunda et al., 2013, Shirakura et al., 2012)

CHAPTER FIVE

RESULTS IN SEBACEOUS GLAND CARCINOMA (SGC)

5.1 Aims of the SGC study

Sebaceous gland carcinoma is an aggressive, potentially fatal cancer that has predilection for the periocular region. Currently, the only treatment is surgical excision and this requires removal of a normal seeing eye to protect life, especially in the pagetoid histological subtype. By understanding the genetic factors, it is hoped that novel therapies can be developed to prevent blinding treatment. A discovery approach using WES of five SGCs was carried out to identify driver mutations, followed by transcriptome analysis with RNAseq. It is hypothesised that its predilection to the periocular region may be associated with UV exposure and a high mutational UV signature would be seen in these tumours. Continuing the transcriptome approach, a histological subtypes analysis was taken to tease out the differences between the nodular (localised) versus pagetoid (spreading) SGC. RNA from 5 pagetoid and 5 nodular subtypes were compared to tarsal plate control tissue and then with each other to differentiate those genes that could be attributed to poor or favourable outcome. Changes at a DNA and RNA level are not the sole changes in cancer, with changes also occurring at an epigenetic and miRNA level that shape the cancer's phenotype. MiRNA from 3 pagetoid were compared to 3 nodular along with 3 normal tarsal plate controls. Any changes were then related back to the mutational and RNA data in an attempt to construct pathway data. Validation of miRNA array data occurred using Taqman real time quantitative PCR. Protein expression was also conducted and an overview of the SGC project is summarised in Figure 5.1.

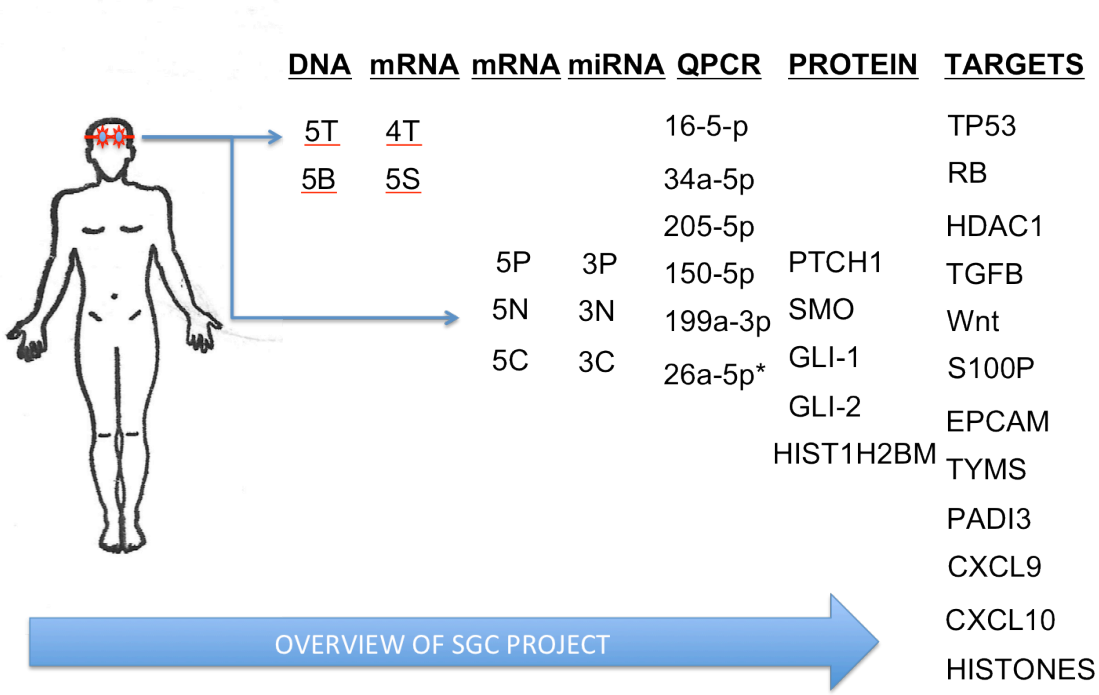


Fig 5.1 SGC Project. Overview of the SGC project highlighting the stepwise progression starting with WES of SGC DNA, RNA sequencing, then different SGC samples undergoing mRNA array, microRNA array and validation using Taqman real time quantitative PCR plus protein expression by immunohistochemistry. Resultant treatment targets are summarised at the end. T, SGC tumour; B, SGC blood match control; S, stromal tissue; P, pagetoid SGC; N, nodular SGC; C, tarsal plate control;*control

4.2 Whole exome sequencing in sebaceous gland carcinoma (SGC)

WES was carried out on five SGC patients and in a similar fashion to the BCC project, posing similar limitations as mentioned in 3.2.

4.2.1 Clinical and histological features of SGC patients

The clinical and histological features are features in Table 5.1 and Figure 5.2.

Sample	Age	M/F	Side	Location/histological subtype	Surgical and adjuvant treatment
SGC1	67	M	L	LUL well differentiated tumour	Localised
SGC2	72	M	R	Invasive and lymph node positive	Exenteration and local radiotherapy
SGC3	69	F	R	RUL Poorly differentiated and in situ disease	Localised
SGC4	57	M	R	Intraepithelial spread	Exenteration
SGC5	61	M	L	Intraepithelial spread	Exenteration

Table 5.1 Demographics of 5 SGC patients who underwent whole-exome and RNA sequencing. Summary of the clinical features for each SGC sample that underwent whole exome and RNA sequencing. M, male; F, female; L, left, R, right; LUL, left upper eyelid; RUL right upper eyelid.

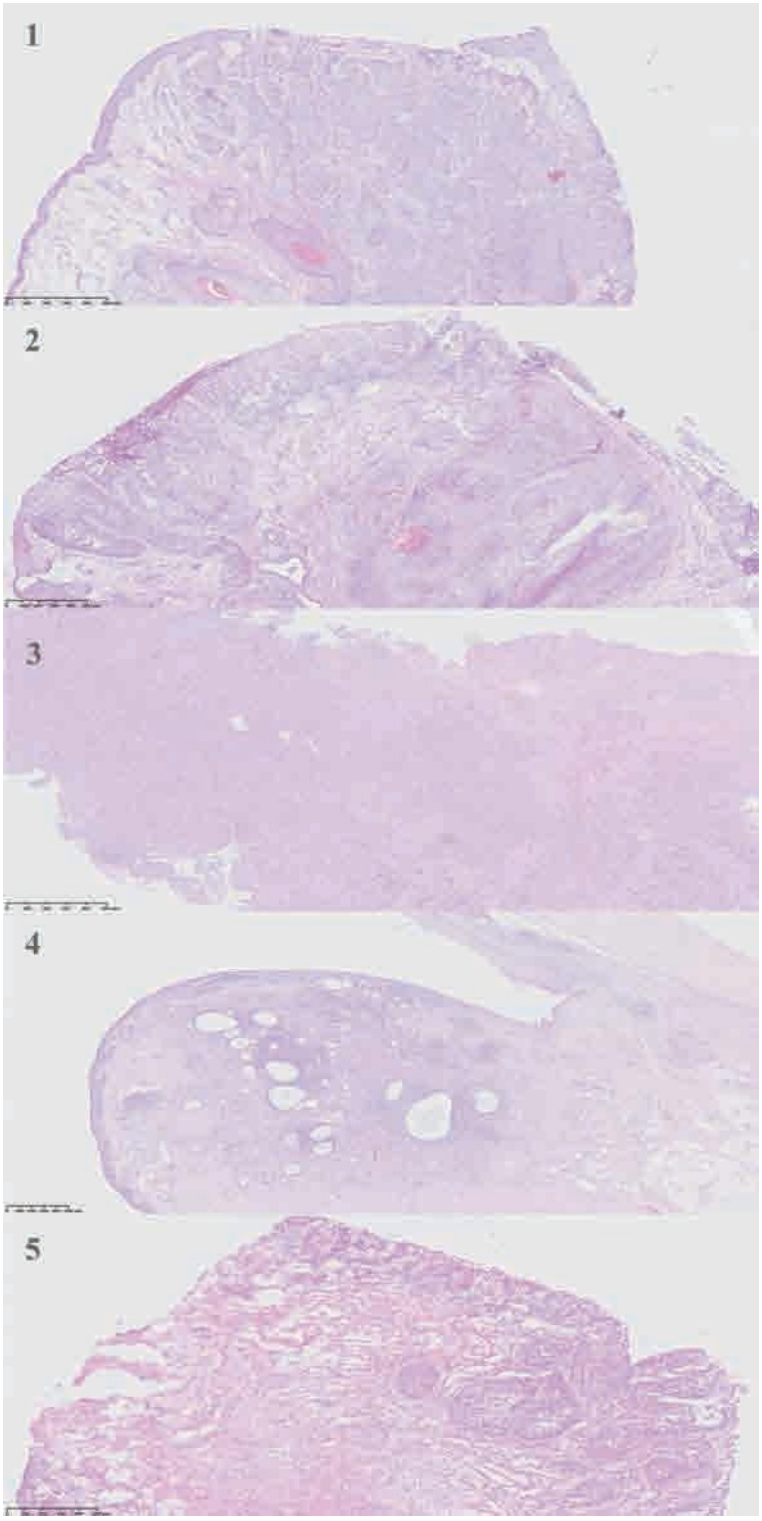


Fig 5.2 SGC samples recruited into SGC Project. Histological collage of SGC samples stained with haematoxylin and eosin (H&E). All samples were initially processed as fresh tissue with fast H&E staining to identify tumour for laser capture. A few cases went on to FFPE processing with further H&E, hence the difference in quality between the samples. Scale bar at 100um.

5.2.2 Read level quality control metrics

Raw WES data is summarised in the tables (Figure 5.3 – 5.6) below without any filtering for depth. Quality control metrics are presented including mean read base quality (Phred score) (Fig 5.4), per sequence quality (Fig 5.5) and GC content per base (Fig 5.6).

	SGC1	SGC2	SGC3	SGC4	SGC5
Number of Variations	71,931	67,937	67,808	64,017	77,131
Variations Overlapping Genes	71,557	67,531	67,416	63,650	76,694
Variations Overlapping Transcripts	71,557	67,531	67,416	63,650	76,694
Variations Overlapping Regulatory Regions	16,360	15,678	15,788	14,348	18,014
Variations Overlapping Protein Domains	68,154	64,013	64,049	60,641	73,072
Intergenic Variations	374	406	392	367	437
Variations With Predicted Serious Consequences	13,267	13,188	13,089	12,109	13,413
Variations With Other Predicted Consequences	58,290	54,343	54,327	51,541	63,281
Non-Synonymous Coding Variations with Consequences	12,497	12,437	12,367	11,436	12,689

Fig 5.3 All variants in SGC. Metrics of all variants identified in WES including those from the single nucleotide polymorphism database (dbSNP) release 138.

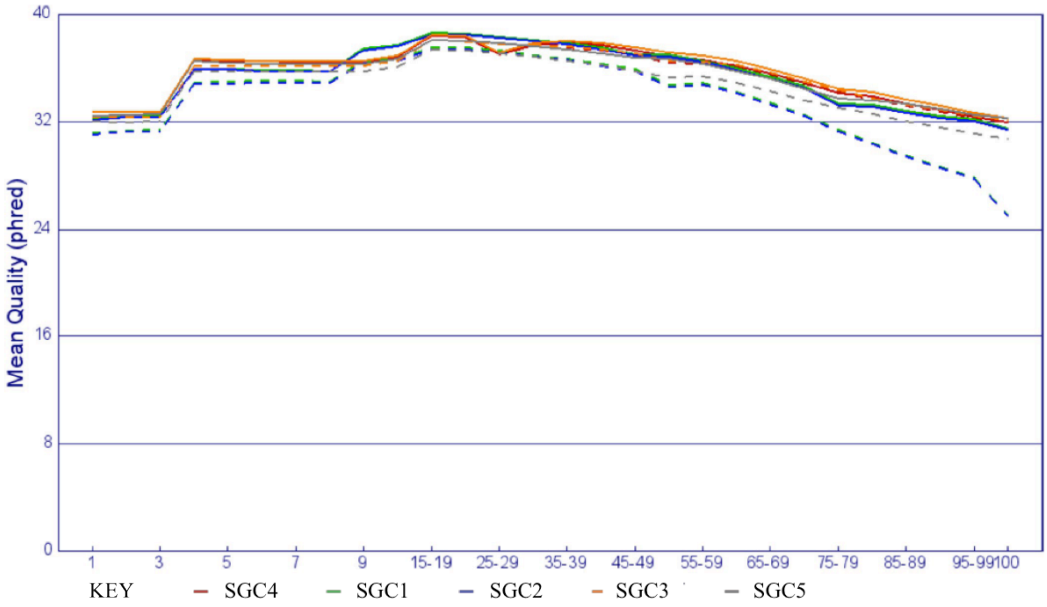


Fig 5.4 Mean Phred score for SGC. Plot of the mean read base quality (Phred score) at each position of a read (cycle) for all five SGC samples. Mean quality decreases towards the end of the read as expected.

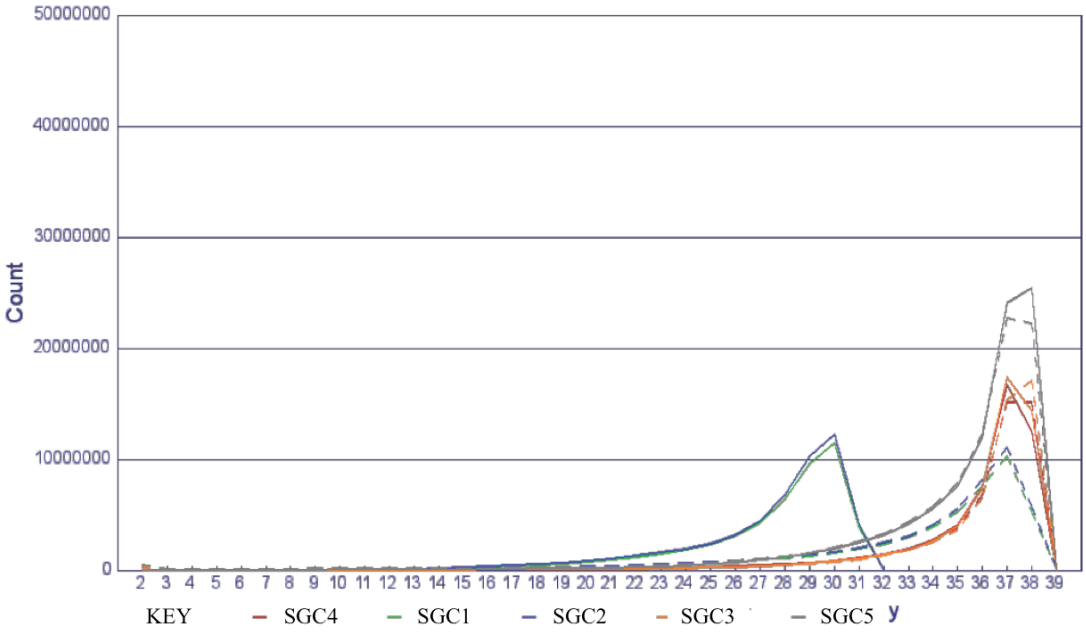


Fig 5.5 Per sequence quality in SGC. Counts of the mean sequence quality for all individual sample's reads. Mean sequence quality is just greater than 30 for all samples, indicating good quality.

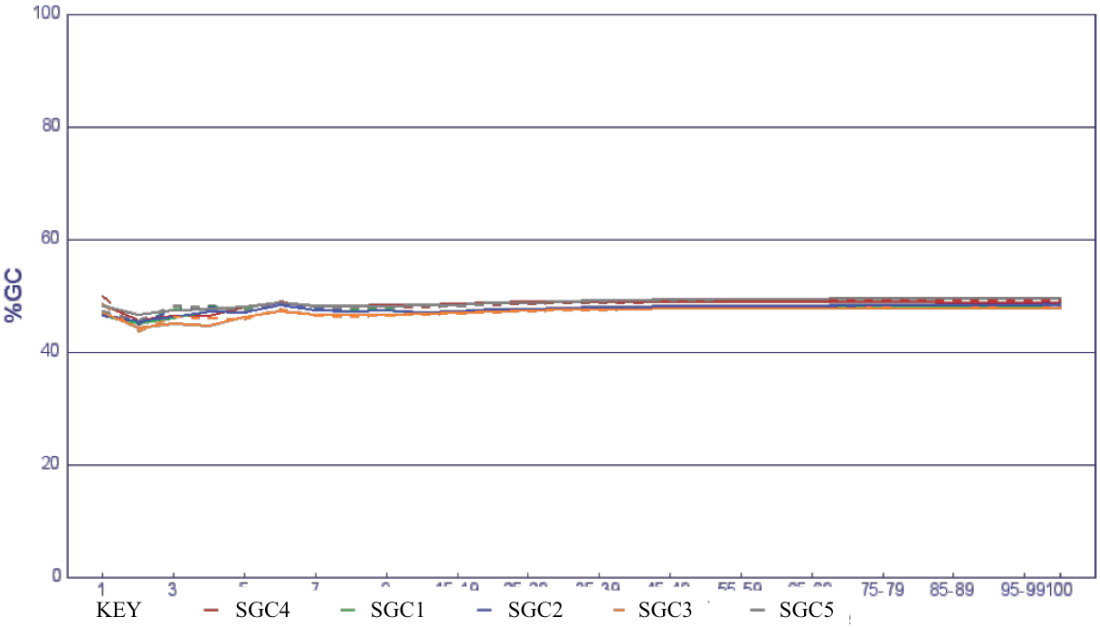


Fig 5.6 Per base GC content for SGC. (A) Plot of percentage GC content at each position of a read (cycle) for all five SGC samples. It reflects the GC content for humans which is between 42-48 percent.

5.2.3 Mutational profile of SGC

The overall mutational burden of potentially pathogenic variants is summarised in Table 5.2. In contrast to BCC, there are a low number of variants reflecting a visceral cancer-like profile rather than a skin tumour. Thus, it seems that SGC is not a tumour arising from skin tissue, but from deeper structures such as the meibomian gland.

Sample	Total	Non-syn	Syn	Ratio of Non-Syn:Syn
RSGC1	50	34	12	2.8
SGC2	615	375	190	2.0
SGC3	265	181	64	2.8
SGC4	259	174	66	2.6
SGC5	296	163	109	1.5

Table 5.2 Mutational burden in SGC. Overall mutational burden, non-synonymous single nucleotide and synonymous single nucleotide variants for each sample that underwent WES after filtering for a minimum of x20 depth. SGC, sebaceous gland carcinoma; Non-syn, non-synonymous single nucleotide variant; syn, synonymous single nucleotide variant.

UV signature and burden

As mentioned in chapter 3, UV signature contains a specific type of base substitution that occurs at dipyrimidine sites, and is a cytosine to thymine (C>T/G>A) change. SGC2 is the only sample with a C>T change at over 90%. At first glance, this may reflect a UV signature in SGC2. The rest are more reflective of a visceral cancer and show only some C>T change. SGC2 is the most aggressive of all five tumours with local lymph node spread and the cancer was present in the sun exposed bulbar conjunctiva along with the tarsal conjunctiva (non-sun exposed). This may explain the larger number of mutations and higher UV burden; at this stage, the control mechanism are all but defunct, thus the usual UV protection is no longer present at the tumour location.

There is a large proportion of G>T (C>A) changes within SGC. This traditionally represents a tobacco signature whereby polycyclic aromatic hydrocarbons are converted to epoxides via liver enzymes which in turn form alkylated guanine adducts that lead to G>T transversion.(Harris, 2013)

Putting the two findings together, G>T (C>A) occurring in the presence of a fair presence of C>T at NpCpG sites is more characteristic of signature 6, a marker of a cancer with defective DNA mismatch repair.(Alexandrov et al., 2013)

The transitions and transversion were also assessed according to VAF to see if the percentage grouping below was reflecting in both low and high VAF. If these changes were mainly enriched for the mutations with low VAF, they could be artefact from oxidative DNA damage during preparation.(Costello et al., 2013) However, the percentages were similar across several VAF categories highlighting these signatures as being more genuine. For example, G>T/C>A changes for SGC3 was 30% at a higher VAF (greater than VAF 0.10) and for SGC4, 90% at the same higher VAF.

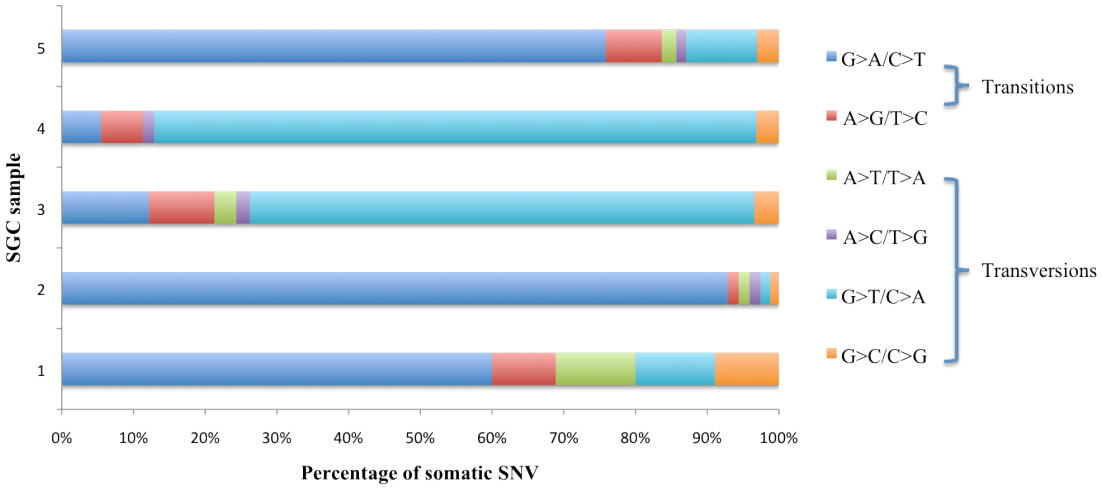


Figure 5.7 Pathological mutations and signature in SGC. Tumour single nucleotide variants are shown for each WES SGC with the percentage of each possible transition/transversion shown.

5.2.4 Driver mutations in periocular SGC

The five SGC samples were analysed together using two different mutational significance algorithms, namely MutSigCV and Intogen, resulting in 14 (Table 5. 4) and 11 (Table 5.5) genes respectively (P <0.05). Both algorithms identified 5 genes in common (Figure 5.8).

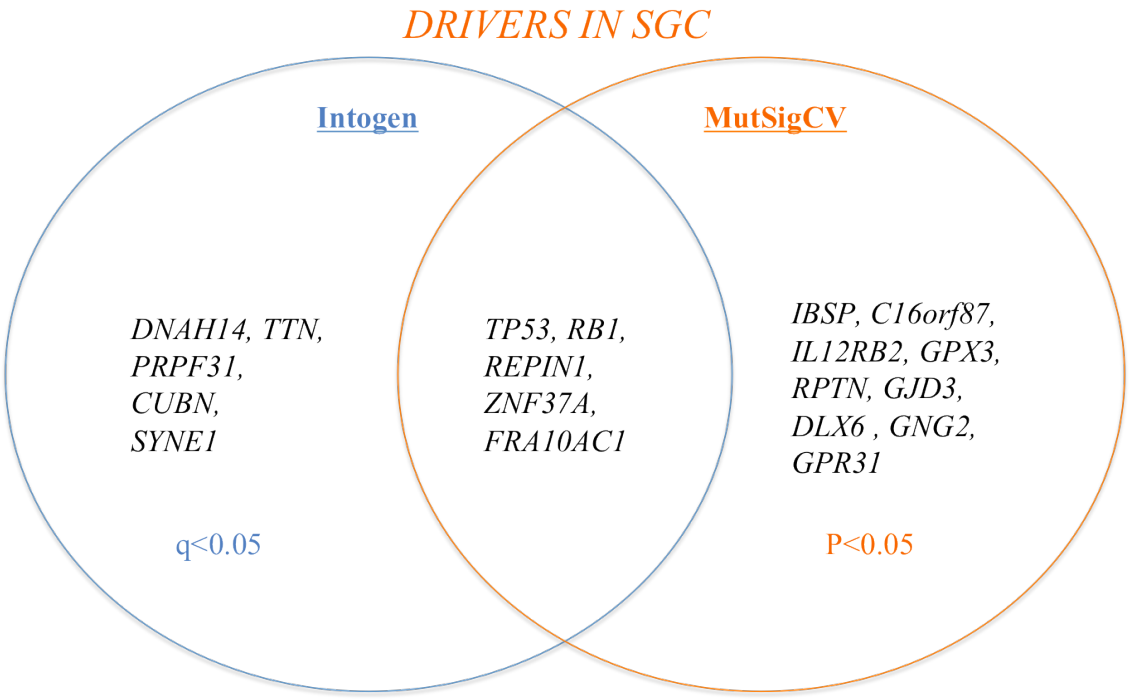


Fig 5.8 Driver genes in SGC. Venn diagram highlighting the significant driver genes in SGC using two algorithms, MutSigCV and Intogen, with five shared genes within the cross over region. q; false discovery rate; p, p-value.

5.2.4.i Shared driver mutations using MutSigCV and Intogen

Fiver genes are shared between the two algorithms, namely TP53, RB1, REPIN1, ZNF37A and FRA10AC1. As mentioned in section 3.3.4, TP53 is the archetypal tumour suppressor gene. It is ubiquitously mutated in all five samples (Table 5.3) with four out of five damaging (in the binding site), but none were G>T (smoking signature). A mutation in TP53 is seen in 3-5% of normal eyelid skin so it may be a passenger in SGC4 where it is predicted to be benign, although the VAF was 0.07 (14% of the sample) so it

may have started to be selected as a driver, despite its benign prediction. (Martincorena et al., 2015)

Sample	DNA	Protein	SIFT	Polypen	COS id/dbSNP
SGC1	c.715_716insTA	p.(Asn239fs)	---	---	
SGC2	c.582_583insAG	p.(Ile194fs)	---	---	
SGC3	c.932G>A	p.Arg248Trp	0, D,H	1, P	10656 / rs121912651
SGC4	c.310C>A	p.Met40Ile	0.22, T,H	0, B	
SGC5	c.582_583insAG	p.(Ile194fs)	---	---	

Table 5.3 *TP53* variants seen in SGC cohort. The DNA location and protein change are highlighted. Fs, non-synonymous frameshift:stop/gain; Analysis from SIFT (D=damaging, T=tolerated, L=low probability, H=high probability); Poly=Polypen (P=probably damaging, B=benign) and Cos id=COSMIC identifier; dbSNP,

5.2.4.ii Drivers identified using MutSigCV algorithm

Gene	NS	S	P Value	Gene	NS	S	P Value
TP53	5	1	6.79E-05	IL12RB2	2	0	2.92E-02
REPIN1	2	2	3.39E-03	GPX3	1	0	3.52E-02
IBSP	2	0	5.30E-03	RPTN	7	1	3.77E-02
FRA10AC1	2	1	6.69E-03	GJD3	1	0	3.94E-02
RB1	2	0	1.13E-02	DLX6	2	0	4.60E-02
ZNF37A	2	0	1.69E-02	GNG2	1	0	4.94E-02
C16orf87	1	0	2.33E-02	GPR31	1	0	4.97E-02

Table 5.4 Drivers in SGC using MutSigCV silent and non-silent mutations. 14 driver genes using MutSigCV taking into account silent and non-silent genes, and ordered according to P value with a P<0.05. S, silent; NS, non-silent.

5.2.4.iii Drivers identified using Intogen algorithm

Gene	Freq	Driver	P-value	Gene	Freq	Driver	P-value
TP53	5	1	0.001	ZNF37A	2	0	0.011
RB1	2	1	0.002	PRPF31	2	0	0.018
REPIN1	2	0	0.002	CUBN	2	0	0.025
DNAH14	2	0	0.002	SYNE1	2	0	0.028
TTN	3	0	0.007	FRA10AC1	2	0	0.050

Table 5.5 Intogen derived drivers in SGC. 10 driver genes arranged using Intogen, p-value < 0.05 and an fm-bias (q-value) value <0.05. 1, known driver gene; 0, driver status unknown; Freq, frequency.

Dynein Axonemal Heavy Chain 14 (*DNAH14*) – chromosome 1q42.12

Heavy chains of axonal dynein are encoded by a family of similar functioning genes including DNAH14. Their exact function is not clear; however, murine studies demonstrate abnormal sister chromatids segregation when mutations occur in the leftright dynein gene. This suggests that DNAH family may be involved in positioning of mitotic spindles, microtubules and chromosomal dynamics. Biallelic mutations of DNAH11 have been seen in primary ciliary dyskinesia.(Knowles et al., 2012) Germline polymorphisms of DNAH14, for example rs1857623 in ovarian cancer, plays a role in overall cancer survival and prognosis.(Braun et al., 2013) DNAH2, 5 and 10 genetic aberrations were associated with renal cell carcinoma as were the spindle assembly and chromosome separation pathways.(Arai et al., 2015) Other members of the dynein family were also mutated (Table 5.6)

Sample	Gene	Chr	Genome / DNA level	Protein level	Change /S P	T Depth	VAF
SGC1	DNAH2	17	c.5838_5839insTTG	p.Phe1941_Gly1942insVal	pepshift	88	0.33
SGC2	DNAH11	7	c.794C>T	p.Gln255*	Stop-gain	128	0.11
SGC3	DNAH14	1	c10903G>T	p.Gly3635*	Stop-gain	151	0.05
	DNAH11	7	c.10002G>A	p.Arg3324His	O,D,H;0.9,P	188	0.26
	DNAH17	17	c.3807C>A	p.Asp1228Tyr	1T,H;0.94 P	43	0.12
SGC4	DNAH14	1	c.12089C>A	p.Ser4030*	Stop-gain	73	0.10
	DNAH7	2	c.7794C>A	p.Glu2565*	Stop-gain	35	0.14
SGC5	DNAH11	2	g.196759699C>T	p.? (Intronic)	---	62	0.34
	DNAH5	5	c.223A>C	p.Leu40Val	0.57 T,H; B	337	0.17

Table 5.6 SGC Dynein gene variants. Non-synonymous changes detected in the Dynein family from all five WES SGC samples and undergoing SIFT (S) and Polypen (P) analysis. Chr, chromosome; AA, amino acid; S, SIFT; P, Polypen; T, tumour; VAF, variant allele frequency.

5.2.5 Periocular SGC WES molecular pathway analysis

Intogen pathway prediction was performed on WES data across the five SGC samples. Fifteen pathways were significantly affected using a p-value < 0.05. This included 7 cancer pathways: pancreatic cancer, chronic myeloid leukaemia, non-small cell lung cancer, melanoma, colorectal cancer, glioma and small cell carcinoma. Using a q-value of <0.10 reduced the number to five pathways.

Identifier	Description	Geneset size	Q-value	Freq
Hsa04110	Cell cycle	124	0.026	5
Hsa04120	Ubiquitin mediated proteolysis	137	0.034	4
Hsa05212	Pancreatic cancer	70	0.091	5
Hsa04310	Wnt signalling pathway	152	0.091	5
Hsa04728	Dopaminergic synapse	130	0.091	5

Table 5.7 Driver pathways in SGC. Intogen pathway prediction highlighting altered pathways using WES data from 5 SGC tumours and using a q value<0.10. Freq, frequency.

5.3 Chromosomal instability in SGC

Single nucleotide changes represent damage to a single strand of DNA and these single base changes underpin mutational instability. Damage to two strands of DNA or unplanned double stranded breaks leads to chromosomal instability with subsequent homologous or non-homologous end joining. Large loss of chromosome regions can inactivate tumour suppressor genes and the short arm of chromosome 17 is lost in SGC1 where TP53 is located (Figure 5.9). Amplification of chromosomal regions can promote tumorigenesis by activation a proto-oncogene and chromosome 8 is significantly amplified where the MYC resides in SGC2 and even more so in SGC1. A large-scale shattering of the chromosomes in a single cellular event is termed chromothripsis whereby the DNA repair machinery joins segments in an erroneous fashion leading to deletions or amplification on a massive scale.

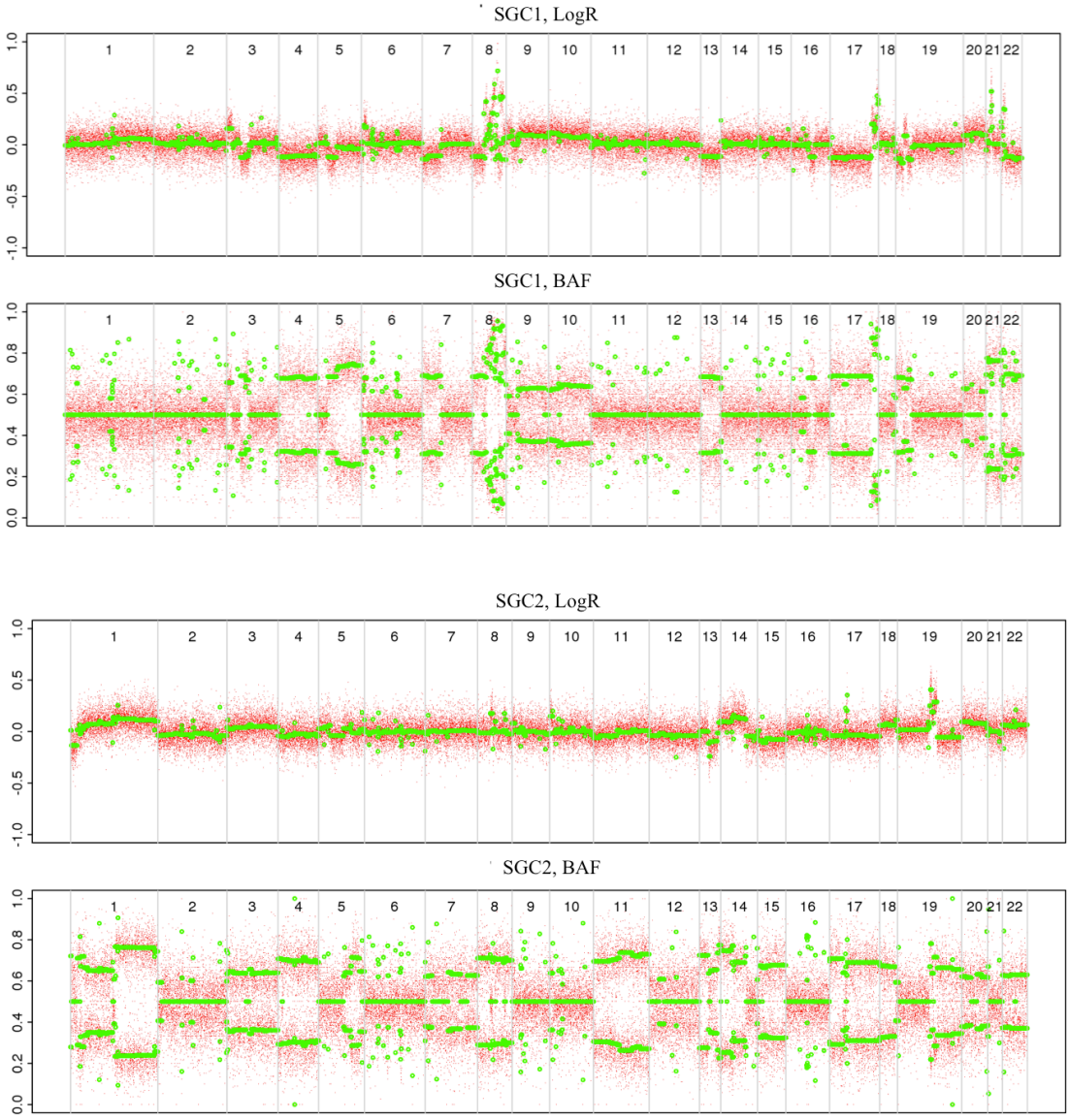


Fig 5.9 Copy number analysis of SGC. Log R Ratio (LogR) and B-allele frequency (BAF) SGC1 and SGC2.

5.4 RNA sequencing in periocular SGC

5.4.1 Clinical features of periocular SGC patients

All five WES samples were sent for RNAseq with normal stroma acting as the reference control (Table 5.1). Two had intraepithelial (pagetoid) spread, two were poorly differentiated tumours with invasion and one was a well-differentiated tumour. The latter underwent local excision and the former three underwent exenteration.

5.4.2 SGC RNA Expression

One tumour sample (SGC4) Phred score tailed off after the 15 cycle and failed quality control, thus was excluded from further analysis. A paired analysis of SGC tumour versus SGC stroma using limma R/Bioconductor software package (voom function) was preformed using raw p value less than 0.01 and log2FC greater than 1 or less than -1.(Ritchie et al., 2015) Downregulation in the tumour is defined as log2FC<-1 and upregulation in the tumour as log2FC > 1. Paired analysis of SGC tumour versus normal stroma revealed 949 differentially expressed genes using a q-value < 0.05 and log2FC >1 or <-1. Log2FC ranged from -6.1 to 5.1 and expression levels from -1.6 to 11.5. A heatmap of the 949 DE findings is summarised in Figure 5.10 and almost 72% (679 genes, blue colour) were downregulated with the vast majority (449) being -2 log2FC or more.

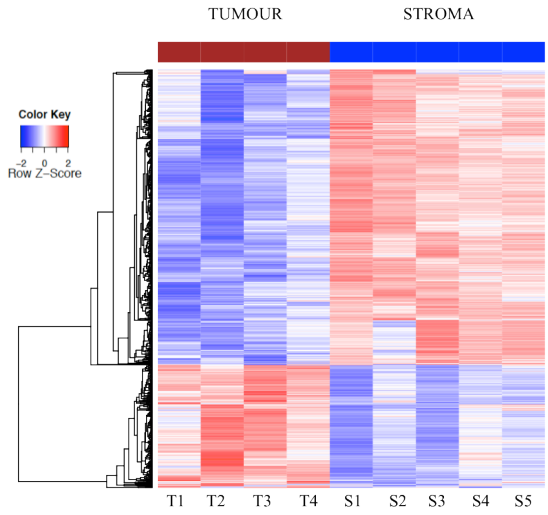


Fig 5.10 Heatmap of SGC tumour and stroma. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression.

5.4.2i Upregulated genes in SGC

Genes found to be significantly upregulated were organised according to their Log2FC (Table 5.8) and then according to the highest expressed (Table 5.9). ON 28% of DE genes were upregulated. Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Log2FC	Exp	P val	Gene	Log2FC	Exp	P val
SPRR2A	4.3	2.9	8.6E-03	S100A8	3.2	9.3	0.03
SERPINB4	3.6	4.5	7.6E-05	S100P	3.1	4.4	0.05
AARD	3.5	2.6	0.02	SERPINB3	2.9	5.6	6.1E-04
SYCP2	3.4	3.8	8.1E-05	HIST1H4H	2.9	2.2	6.7E-03
SPRR2D	3.3	3.1	2.7E-04	EPCAM	2.9	6.6	4.8E-03

Table 5.8 Highest Log2FC upregulated genes in SGC. Top 10 differentially expressed RNA are arranged according to the highest log2FC with an average expression greater than 2.0 and P <0.05. Exp=expression, P-val=P-value.

Gene	Log2FC	Exp	P val	Gene	Log2FC	Exp	P val
S100A8	3.2	9.3	0.03	MGST1	2.3	8.4	2.14E-05
FABP5	2.2	9.3	0.001	KRT16	2.2	8.0	7.2E-03
STMN1	2.1	9.0	8.1E-04	S100A7	2.8	7.7	9.2E-03
FXYD3	2.8	8.6	2.7E-03	CHCHD7	2.1	7.7	0.01
S100A2	2.2	8.5	0.04	EPCAM	2.9	6.6	4.8E-03

Table 5.9 Highest expressed genes in SGC. Top 10 differentially, highest expressed, upregulated RNA with a log2FC of greater than 2.0 and P<0.05. Exp=expression, P-val=P-value.

5.4.2.ii Downregulated genes in SGC

As mentioned, the majority of genes are downregulated in SGC. Genes found to be significantly downregulated were organised according to their Log2FC (Table 5.10) and then according to the highest expressed (Table 5.11). Selected genes of interest are further discussed in terms of their biological relevance and cancer association.

Gene	Log2FC	Exp	P-val	Gene	Log2FC	Exp	P-val
PYGM	-6.1	5.6	1.6E-03	CHGB	-5.3	4.8	4.1E-03
MYH4	-5.6	3.7	3.9E-03	RP11-755O11.2	-5.3	2.9	4.3E-03
PVALB	-5.5	3.5	1.4E-03	SRL	-5.2	3.4	8.5E-04
ATP2A1	-5.5	6.6	2.6E-03	CKM	-5.2	8.3	0.02
MYH1	-5.4	6.6	9.4E-03	CASQ1	-5.2	5.8	2.2E-03

Table 5.10 Downregulated genes with highest Log2FC in SGC. Top 10 differentially expressed RNA are arranged according to the lowest log2FC with an average expression greater than 2.0. Exp=expression, P-val=P-value.

Gene	Log2FC	Exp	P-val	Gene	Log2FC	Exp	P-val
IGHG1	-2.5	11	0.01	COL3A1	-2.3	9.6	4.7E-03
CD74	-2.5	10	0.01	TTN	-3.4	9.4	0.05
GSN	-2.3	10	4.3E-05	APOE	-2.2	9.4	0.03
COL1A2	-2.1	10	6.3E-03	COL1A1	-2.1	9.4	0.02
TPM2	-3.7	9.6	1.4E-03	DCN	-2.8	9.2	1.3E-03

Table 5.11 Highest expressed, down regulated genes in SGC. Top 10 differentially, highest expressed, downregulated RNA with a log2FC of less than -2.0. Exp=expression, P-val=P-value.

5.4.3 RNAseq pathway analysis in SGC

RNA sequencing expression was assessed using GSEA to identify significantly altered genesets. Pathway expression was first assessed on an individual tumour basis by comparing tumour with the stroma followed by a comparison between the two subtypes. Significant pathway alteration was determined using a FDR q-value ≤ 0.10.

5.4.3.i SGC expression using KEGG genesets

Twenty-seven pathways were shown to be upregulated when analysing RNAseq gene-set data (Table 5.12). Upregulation occurred in 4 DNA repair pathways; mismatch repair, homologous recombination, nucleotide excision repair and base excision repair.

KEGG pathway	NES	GS Size	Q-val	P-val
Ribosome	4.0	86	<0.001	<0.001
Spliceosome	3.2	125	<0.001	<0.001
DNA_replication	3.0	36	<0.001	<0.001
Cell_cycle	2.8	121	<0.001	<0.001
Proteasome	2.7	41	<0.001	<0.001
Oxidative_phosphorylation	2.6	107	<0.001	<0.001
Mismatch_repair	2.5	23	<0.001	<0.001
Pyrimidine_metabolism	2.4	89	<0.001	<0.001
RNA_degradation	2.4	56	<0.001	<0.001
Nucleotide_excision_repair	2.4	43	<0.001	<0.001
Homologous_recombination	2.2	28	<0.001	<0.001
Protein_export	2.2	24	<0.001	<0.001
Terpenoid_backbone_biosynthesis	2.2	15	<0.001	<0.001
Rna_polymerase	2.1	28	1.1E-04	<0.001
Huntingtons_disease	2.0	156	8.9E-04	<0.001
Steroid_biosynthesis	2.0	15	0.001	<0.001
Parkinsons_disease	2.0	103	0.001	<0.001
Basal_transcription_factors	1.9	30	0.002	<0.001
Butanoate_metabolism	1.8	27	0.005	<0.001
Glycosylphosphatidylinositol_GPI_anchor_bisynthesis	1.8	24	0.007	0.018
Base_excision_repair	1.8	33	0.008	0.004
Oocyte_meiosis	1.7	103	0.011	<0.001
Systemic_lupus_erythematosus	1.7	78	0.018	0.005
P53_signaling_pathway	1.6	67	0.019	0.005
One_carbon_pool_by_folate	1.6	16	0.021	0.028
Alzheimers_disease	1.6	147	0.021	<0.001
Peroxisome	1.5	74	0.047	0.0262

Table 5.12 Upregulated SGC pathways using GSEA on KEGG genesets. Upregulated KEGG pathway expression in SGC tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.05. NES, normalised enrichment score; Q-value, false discovery rate; P-val, P-value

More pathways were downregulated (Table 5.13) with the most downregulated in extracellular matrix receptor interaction and 2 others involved in cell adhesion; focal adhesion and cell adhesion molecules. This may contribute to its ability to spread locally, especially in the pagetoid subtype.

KEGG pathway	NES	GSS	Q-val	P-val
ECM_receptor_interaction	-2.3	75	<0.001	<0.001
Dilated_cardiomyopathy	-2.3	79	<0.001	<0.001
Focal_adhesion	-2.2	191	<0.001	<0.001
Viral_myocarditis	-2.2	67	<0.001	<0.001
Hypertrophic_cardiomyopathy_HCM	-2.2	74	<0.001	<0.001
Vascular_smooth_muscle_contraction	-2.2	98	<0.001	<0.001
Cell_adhesion_molecules_cams	-2.2	121	<0.001	<0.001
Complement_and_coagulation_cascades	-2.1	48	<0.001	<0.001
Calcium_signaling_pathway	-2.1	139	<0.001	<0.001
JAK_STAT_signaling_pathway	-2.1	107	<0.001	<0.001
Hematopoietic_cell_lineage	-2.1	69	<0.001	<0.001
Intestinal_immune_network_for_IGA_production	-2.0	38	8.3E-05	<0.001
MAPK_signaling_pathway	-2.0	230	1.5E-04	<0.001
Graft_versus_host_disease	-2.0	29	2.8E-04	<0.001
Autoimmune_thyroid_disease	-2.0	31	3.3E-04	<0.001
Cytokine_cytokine_receptor_interaction	-2.0	194	4.2E-04	<0.001
Axon_guidance	-2.0	119	4.5E-04	<0.001
Arrhythmogenic_right_ventricular_cardiomyopathy_ARVC	-1.9	66	4.8E-04	<0.001
Lysosome	-1.9	116	0.001	<0.001
Allograft_rejection	-1.9	29	0.001	0.002
Inositol_phosphate_metabolism	-1.9	52	0.002	0.001
Leishmania_infection	-1.8	64	0.003	<0.001
Glycosaminoglycan_degradation	-1.8	19	0.003	0.003
Chemokine_signaling_pathway	-1.8	164	0.004	<0.001
Neuroactive_ligand_receptor_interaction	-1.8	137	0.004	<0.001
Glycosaminoglycan_biosynthesis_chondroitin_sulfate	-1.8	21	0.005	0.005
Long_term_depression	-1.8	57	0.005	<0.001
T_cell_receptor_signaling_pathway	-1.8	97	0.005	<0.001
Phosphatidylinositol_signaling_system	-1.8	73	0.006	0.002

Table 5.13 Downregulated SGC pathways using GSEA on KEGG genesets. Downregulated KEGG pathway expression in SGC tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.006. NES, normalised enrichment score; GSS, Geneset size; Q-value, false discovery rate; P-val, P-value.

5.4.3.ii SGC expression using REACTOME genesets

A second algorithm, REACTOME, was used to assess SGC pathway. Upregulated pathways mainly involve replication (Table 5.14). Downregulated pathways using REACTOME also demonstrated reduction in extracellular matrix pathways (Table 5.15)

REACTOME pathway	NES	GSS	Q-val	P-val
Influenza_viral_RNA_transcription_and_replication	3.8	101	<0.001	<0.001
SRP_dependent_cotranslational_protein_targeting_to_membrane	3.8	109	<0.001	<0.001
Influenza_life_cycle	3.8	135	<0.001	<0.001
Peptide_chain_elongation	3.8	85	<0.001	<0.001
Translation	3.8	147	<0.001	<0.001
3_UTR_mediated_translational_regulation	3.7	106	<0.001	<0.001
Nonsense_mediated_decay_enhanced_by_the_exon_junction_complex	3.7	106	<0.001	<0.001
Metabolism_of_mRNA	3.6	207	<0.001	<0.001
Metabolism_of_RNA	3.6	252	<0.001	<0.001
Mitotic_M_M_G1_phases	3.5	163	<0.001	<0.001
DNA_replication	3.5	183	<0.001	<0.001
Formation_of_ternary_complex_and_subsequently_the_43s_complex	3.4	49	<0.001	<0.001
Cell_cycle_mitotic	3.3	303	<0.001	<0.001
Cell_cycle_checkpoints	3.3	110	<0.001	<0.001
S_phase	3.3	105	<0.001	<0.001
G1_S_transition	3.3	106	<0.001	<0.001
Activation_of_the_mrna_upon_binding_of_the_cap_binding_complex	3.3	57	<0.001	<0.001
Synthesis_of_DNA	3.3	89	<0.001	<0.001
Cell_cycle	3.3	364	<0.001	<0.001
G1_G1_S_phases	3.3	130	<0.001	<0.001
M_G1_transition	3.2	77	<0.001	<0.001
Telomere_maintenance	3.1	48	<0.001	<0.001

Table 5.14 Upregulated SGC pathways using GSEA on REACTOME genesets. Upregulated REACTOME pathway expression in SGC tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.01. NES, normalised enrichment score; GSS, Geneset size; Q-value, false discovery rate; P-val, P-value.

REACTOME pathway	NES	GS Size	Q-val	P-val
Striated_muscle_contraction	-2.4	26	<0.001	<0.001
Muscle_contraction	-2.3	45	<0.001	<0.001
Extracellular_matrix_organization	-2.3	77	<0.001	<0.001
Integrin_cell_surface_interactions	-2.2	73	1.8E-04	<0.001
NCAM1_interactions	-2.2	35	1.4E-04	<0.001
Collagen_formation	-2.1	56	1.2E-04	<0.001
Formation_of_fibrin_clot_clotting_cascade	-2.1	15	1.0E-04	<0.001
Glycosaminoglycan_metabolism	-2.1	96	1.9E-04	<0.001
Nitric_oxide_stimulates_guanylate_cyclase	-2.0	24	8.4E-04	<0.001
Gpcr_downstream_signaling	-2.0	313	0.001	<0.001
Degradation_of_the_extracellular_matrix	-2.0	21	0.001	<0.001
Myogenesis	-2.0	24	9.9E-04	<0.001
Signaling_by_PDGF	-2.0	113	0.00144768	<0.001
Platelet_homeostasis	-2.0	67	0.001	<0.001
Axon_guidance	-2.0	223	0.002	<0.001
Chondroitin_sulfate_dermatan_sulfate_metabolism	-1.9	45	0.002	<0.001
Platelet_activation_signaling_and_aggregation	-1.9	171	0.004	<0.001
Signaling_by_RHO_gtpases	-1.9	109	0.004	<0.001
NCAM_signaling_for_neurite_out_growth	-1.9	59	0.003	<0.001
Phospholipase_C_mediated_cascade	-1.9	42	0.006	0.001
Signaling_by_GPCR	-1.9	396	0.006	<0.001
G_alpha_S_signalling_events	-1.9	74	0.006	<0.001
IL_3_5_and_gm_csf_signaling	-1.8	38	0.007	<0.001
Nuclear_receptor_transcription_pathway	-1.8	38	0.007	<0.001
Smooth_muscle_contraction	-1.8	23	0.008	0.001
Synthesis_of_pips_at_the_plasma_membrane	-1.8	29	0.008	0.001
Developmental_biology	-1.8	338	0.01	<0.001

Table 5.15 Downregulated SGC pathways using GSEA on REACTOME genesets. Downregulated REACTOME pathway expression in SGC tumour versus stroma. Arranged according to NES with a false discovery rate (q-value) <0.006. MSigDB, molecular signature database; NES, normalised enrichment score; GSS, Geneset size; Q-value, false discovery rate; P-val, P-value.

5.5 Correlated SGC WES and RNA sequencing data

5.5.1 Individual driver genes with RNA changes in SGC

Driver genes identified using WES data were assessed in terms of their expression. SGC driver genes (see section 5.2.4) are listed in Table 5.16. Retinoblastoma gene was downregulated, as were REPIN1, CUBN, SYNE1, GPX3 and GNG2. Individual TP53 showed a possible trend to downregulation, but at a pathway level is significantly upregulated (Table 5.12). DNAH14 is upregulated.

Driver	Log2FC	Exp	P-val	Driver	Log2FC	Exp	P val
TP53	-0.2	5.0	0.598	IBSP	---	---	---
RB1	-0.9	5.2	0.010	C16orf87	1.0	4.3	0.041
REPIN1	-0.7	6.0	0.024	IL12RB2	0.4	2.0	0.661
ZNF37A	-0.5	4.8	0.239	GPX3	-2.0	7.2	<0.001
FRA10AC1	0.2	4.7	0.588	RPTN	---	---	---
DNAH14	2.0	3.1	0.002	GJD3	-1.3	0.8	0.08
TTN	-3.4	9.4	0.046	DLX6	2.0	0.2	0.13
PRPF31	0.01	6.4	0.966	GNG2	-2.6	4.4	0.005
CUBN	-1.6	1.1	0.011	GPR31	---	---	---
SYNE1	-2.3	5.4	0.002				

Table 5.16 Driver gene expression in SGC. Shared and individual driver genes selected by Intogen and MutSigCV with a p-value <0.05. Exp, expression; P-val, P-value.

5.5.2 Pathway link of SGC DNA and RNA analysis

In a similar fashion to the driver genes, significantly modified pathways detected using WES data was compared to pathways picked up using RNAseq. Comparison for SGC data is summarised in Table 5.17. Upregulation of the cell cycle and ubiquitin-mediated proteolysis was seen at a DNA and RNA level, as was Wnt signalling. Downregulation of the melanoma pathway is seen, but of unknown significance.

	WES				**RNAseq**				
Identifier	**Description**	**GS**	**q-val**	**F**	**NES**	**GS**	**q-val**	**P-val**	**R/K**
Hsa04110	Cell cycle	124	0.026	5	2.8	121	<0.001	<0.001	K
Hsa04120	Ubiquitin mediated proteolysis	137	0.034	4	2.6	46	0	0	R
Hsa05212	Pancreatic cancer	70	0.091	5	-	-	-	-	-
Hsa04310	Wnt signalling pathway	152	0.091	5	2.5	61	0	0	R
Hsa04728	Dopaminergic synapse	130	0.091	5	-	-	-	-	-
Hsa05218	Melanoma	71	0.13	5	-1.7	59	0.009	0.001	K

Table 5.17 Correlation of SGC WES and RNAseq pathway analysis. Intogen pathway prediction highlighting altered pathways using WES data from 5 SGC tumours. Pathway expression when comparing SGC tumour versus stroma utilising both KEGG and REACTOME genesets. Kegg ID, KEGG pathway identification number; q-val, q-value; Fr, frequency of mutations within tumour sample; NES, normalised enrichment score; GS, geneset size; p-val, P-value; K, KEGG geneset; R, REACTOME geneset.

5.6 SGC RNA subtype analysis

A second method of RNA analysis was performed on a separate, larger cohort of SGC patients whose details are below. These are labelled sebaceous cell carcinoma (SCC) rather than SGC to avoid confusion between tissue samples used for RNA validation and miRNA analysis, and the five used for sequencing. This served two purposes: firstly, to carry out a histological subtype analysis to decipher any differences between nodular SGC and the more aggressive pagetoid subtype. Secondly, it provides a source of validation of the RNAseq data by identifying any common genes that are differentially expressed using both modalities. RNA expression from 4 nodular (SCC1-4) and 4 pagetoid (SCC 7,9-11) SGC samples were compared to 4 normal tarsal plate control with a subsequent cross comparison. A differential expression (DE) analysis using a threshold of $P < 0.01$ identified 1625 DE probe sets (Fig 5.11), 590 in nodular, 796 in pagetoid and 239 shared and visually represented as a heatmap (Fig 5.12). Using a threshold of false discovery of < 0.05, it narrowed the search to 9 nodular, 9 pagetoid and 4 shared genes. (Fig 5.13)

Microarray analysis was performed to assess the RNA and microRNA (section 5.7) of the two histological subtypes. Probes on an array have the problem that they often detect more than one gene. The Affymetrix and Nanostring array chips used employ short oligonucleotides to reduce the likely of detecting more than one gene, however, it is still possible. Some of the array oligonucleotides are transcript specific, but some primary transcripts undergo alternative splicing depending on the cell status and these alternative transcripts can have opposing mechanism of action. Furthermore, closely related genes may cross hybridise with a probe despite having the ‘perfect match oligonucleotide’. Moreover, tumours contain novel splice variants and duplicated genes that either may not be detected or can skew the hybridisation and subsequent fluorescence.

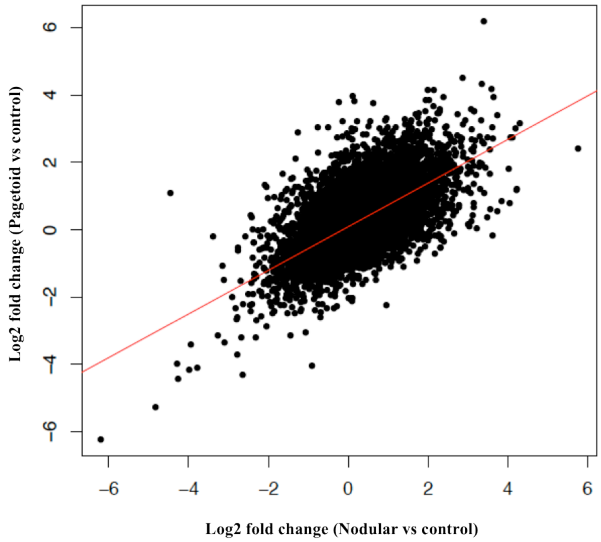


Fig 5.11 Differential expression of RNA in Pagetoid versus Nodular SGC. Log2 fold change between (pagetoid vs. control) and (nodular vs. control) for all 21,446 filtered probe sets.

5.6.1 Clinical and histological features of SCC patients

A different set of SGC patients underwent RNA subtype analysis and coded as SCC rather than SGC to differentiate from the WES/RNAseq cohort (Table 5.18).

Sample	Age	Gender	L	Subtype
SCC1	70	F	RE	Pagetoid
SCC2	71	F	LUL	Pagetoid
SCC3	64	F	RE	Pagetoid
SCC4	66	M	RE	Pagetoid
SCC6	59	F	RUL	Pagetoid
Sample	**Age**	**Gender**	**L**	**Subtype**
SCC7	66	F	RUL	Nodular
SCC8	72	F	LLL	Nodular
SCC9	78	M	LUL	Nodular
SCC10	65	F	LUL	Nodular
SCC11	75	F	RLL	Nodular

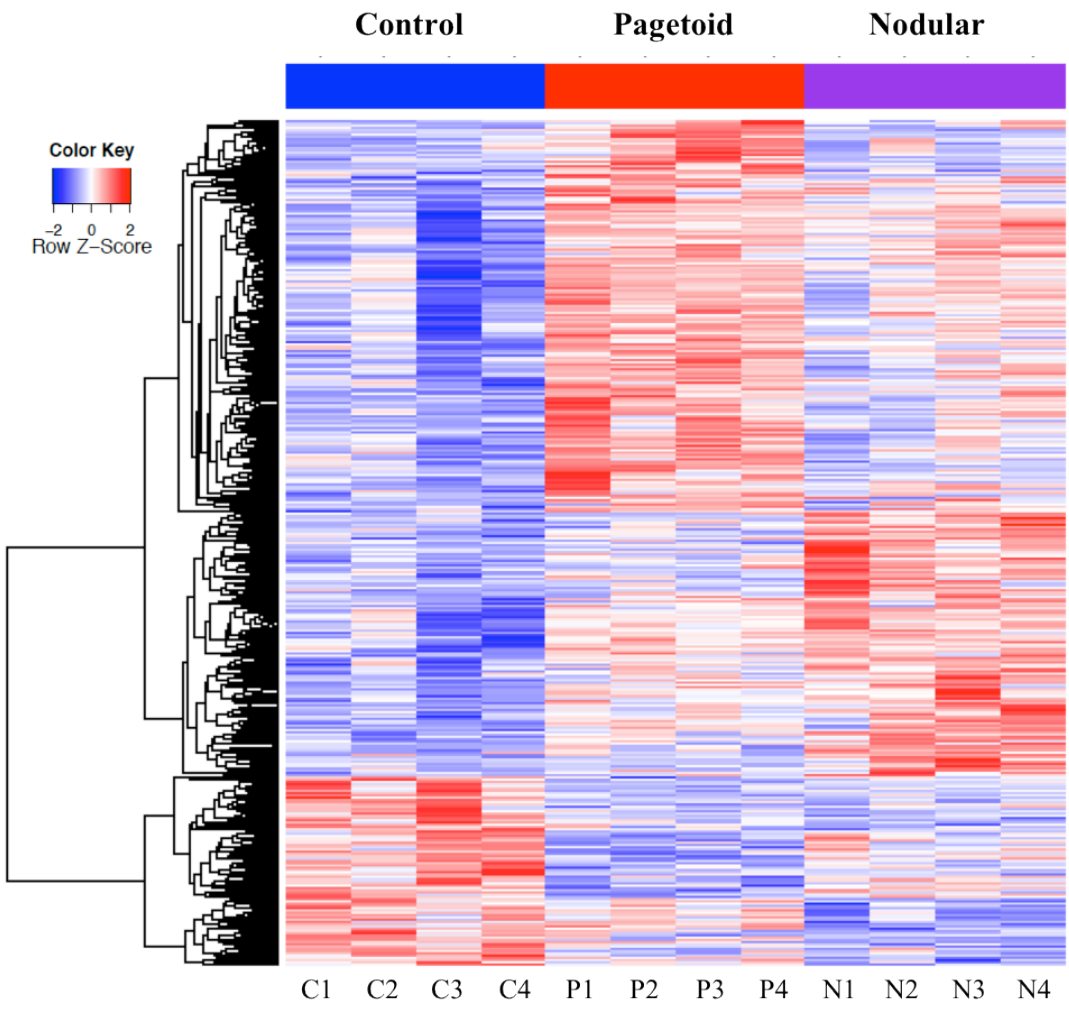


Fig 5.12 Heatmap of RNA in Tarsal plate control, Pagetoid SGC and Nodular SGC. Log2 fold change between (pagetoid vs. control) and (nodular vs. control) for all 1625 differentially expressed probe sets p <0.01. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression. C, tarsal plate control; P, pagetoid SGC; N, nodular SGC.

5.6.2 Nodular SGC messenger RNA

Nodular SGC had 590 DE genes and the top nodular specific are listed in Table 5.19. Selected individual genes are discussed in more detail below.

	Nodular vs. Control			Pagetoid vs. Control		
Gene	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P val**
RNU7-141P	5.8	8.4	<0.001	2.4	5.1	0.001
Y_RNA	3.6	6.6	<0.001	2.7	5.6	0.001
SYCP2	3.3	6.2	<0.001	2.0	4.9	0.002
SNORD51	3.2	6.4	<0.001	0.7	3.9	0.124
SCARNA8	2.8	7.5	<0.001	0.7	5.4	0.144
RNU7-57P	2.4	5.5	<0.001	0.6	3.7	0.113
SCARNA1	2.4	5.8	<0.001	1.2	2.0s	0.005

Table 5.19 Nodular specific RNA. Differentially expressed genes present in nodular SGC only when normalised to tarsal plate control and using a false discovery rate <0.05 and Log2FC fold change >1.5. Pagetoid readings included to highlight the difference between subtypes. LC, Log2FC; Exp, average expression; P-val, p-value.

.

5.6.3 Pagetoid SGC messenger RNA

Pagetoid SGC demonstrated 796 DE genes with the top pagetoid specific listed in Table 5.20. Individual genes are explored below in more detail.

	Pagetoid vs. Control			Nodular vs. Control		
Gene	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P-val**
IGKJ3	6.2	10	<0.001	3.4	7.2	0.005
CXCL9	3.8	7.3	<0.001	0.2	3.6	0.789
IGKV1-5	3.8	6.0	<0.001	-0.2	2.0	0.635
CXCR4	3.0	8.3	<0.001	-0.5	4.7	0.344
DBF4P1	3.0	6.2	<0.001	1.6	4.8	0.006
RP11-44K6.2	3.0	4.5	<0.001	0.2	1.7	0.727
HIST1H2BM	2.8	8.0	<0.001	1.7	6.9	0.003
CKS2	2.8	6.1	<0.001	1.4	4.6	0.002
NUSAP1	2.3	5.7	<0.001	1.9	5.3	<0.001
C2orf47	2.0	5.3	<0.001	0.9	4.3	0.015
CDKN3	2.0	4.1	<0.001	1.6	4.4	<0.001
SMCHD1	1.8	6.5	<0.001	1.0	7.2	0.003

Table 5.20 Pagetoid specific RNA. Differentially expressed genes present in pagetoid SGC only when normalised to tarsal plate control and using a false discovery rate <0.05 and Log2FC fold change >1.5. Nodular expression is included in order to highlight the difference between subtypes. LC, Log2FC; Exp, average expression; P-val, p-value.

5.6.4 Comparison of pagetoid and nodular SGC

Shared genes between the two subtypes numbered 239 using P < 0.01, however, this was reduced to 22 DE genes using a q-value <0.05 (Fig 5.13A). Of these 4 were common to both (Fig 5.14B) and listed in Table 5.21.

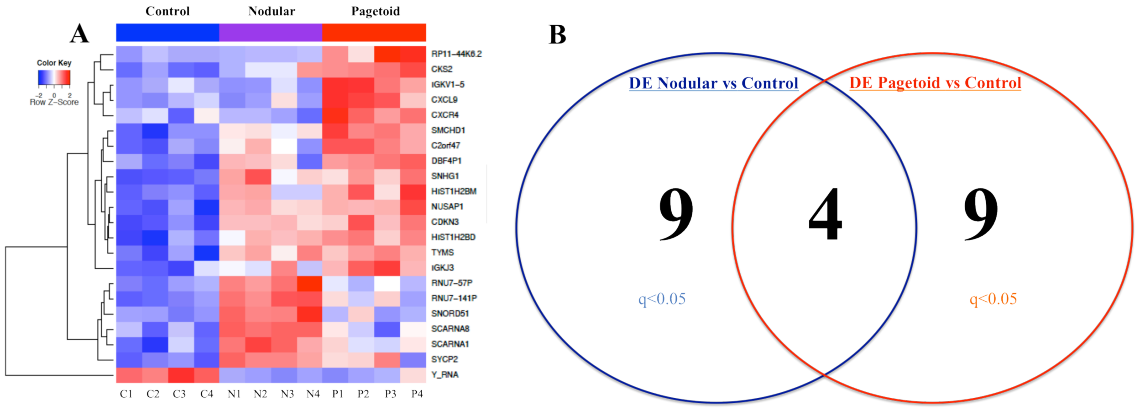


Fig 5.13 Heatmap of shared RNA between pagetoid SGC and nodular SGC. (A) Log2 fold change between (pagetoid vs control) and (nodular vs control) for all 22 differentially expressed genes using a threshold of false discovery rate <0.05. Genes and samples were clustered based on Pearson's correlation. The scaled expression of each gene, denoted as the row Z-score, is plotted in red–blue colour scale with red indicating high expression and blue indicating low expression. (B) Venn diagram highlighting only four genes are shared between nodular and pagetoid subtype .DE, differential expression; C, tarsal plate control; P, pagetoid SGC; N, nodular SGC.

	Nodular vs. Control			Pagetoid vs. Control		
Gene	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P-val**
TYMS	2.6	6.8	<0.001	2.8	7.0	<0.001
HIST1H2BD	2.4	6.4	<0.001	3.2	7.0	<0.01
SNHG1	2.4	6.0	<0.001	2.5	6.0	<0.001
Y_RNA	-3.9	2.8	<0.001	-3.4	3.3	<0.001

Table 5.21 Shared nodular and pagetoid SGC RNA. Differentially expressed genes present in both nodular and pagetoid SGC when normalised to tarsal plate control and using a false discovery rate <0.05 and Log2FC fold change >1.5. LC, Log2FC; Exp, average expression; P-val, p-value.

5.7 SGC Micro RNA subtype analysis

MiRNA expression in three This had led to the potential for immunotherapy and the development of vaccines against the antigen.(Gunda et al., 2013, Shirakura et al., 2012) pagetoid (SCC1,2,3) were compared to three nodular (SCC7,8,9) to decipher the differences between the two subtypes.

5.7.1 Clinical and histological features of SGC patients

Six of the RNA microarray patients also went on to have their miRNA analysed (Table 5.22).

Sample	**Age**	**Gender**	**L**	**Subtype**
SCC1	70	F	RE	Pagetoid
SCC2	71	F	LUL	Pagetoid
SCC3	64	F	RE	Pagetoid
Sample	**Age**	**Gender**	**L**	**Subtype**
SCC7	66	F	RUL	Nodular
SCC8	72	F	LLL	Nodular
SCC9	78	M	LUL	Nodular

Table 5.22 Demographics of SGC samples undergoing miRNA analysis. SCC, sebaceous gland carcinoma; F, female; M, male; L, main location of tumour; R, right; L, left; UL, upper eyelid; LL, lower eyelid.

5.7.2 Nodular SGC microRNA

Seventy-five differentially expressed genes were found to be unique to nodular SGC compared to control and pagetoid using a P value <0.05 (Fig 5.14). The majority of these were downregulated with 27 up regulated. Log2FC ranged from -4.8 to 3.0. Expression ranged from 0.9 to 9.9. The top 5 nodular specific miRNA are listed in Table 5.23 and selected miRNA discussed further below.

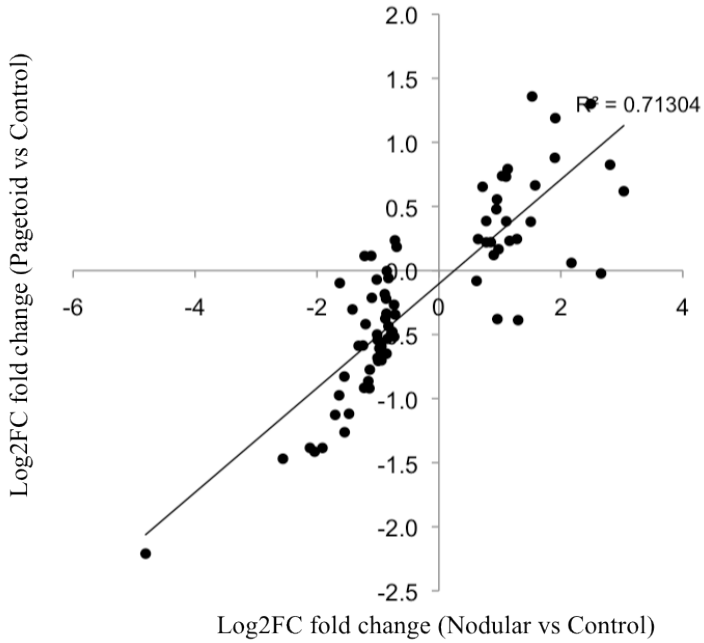


Fig 5.14 Nodular SGC specific microRNA. Seventy-five significant differentially expressed microRNA unique to nodular SGC using a p <0.05 threshold.

	Nodular vs. Control			Pagetoid vs. Control			
miRNA	**LC**	**Exp**	**P-Val**	**LC**	**Exp**	**P-Val**	**Gene Target**
150	3.0	6.4	0.018	0.62	6.4	0.572	ZEB1 in SGC; MMP14; P2X7; HER2; MUC4
142-3p	2.8	9.1	<0.001	0.8	9.0	0.139	FZD7; TGFBR1; PROM1; ABCG2; LRG5; FZD7
143	2.2	5.6	0.028	0.06	5.6	0.945	GABARAPL1; HX2
548g	-1.6	7.3	0.010	-0.09	7.3	0.849	FHIT
603	-2.6	8.0	0.007	-1.5	8.0	0.075	CTNNBIP1; WIF1; E2F1; CCND1; CCND2

Table 5.23 Nodular SGC specific microRNA. Top 5 differentially expressed genes present in nodular SGC only when normalised to tarsal plate control and using a p-value <0.05 and Log2FC fold change >1.5. Pagetoid expression shown to demonstrate non-significant change within the subtype. LC, Log2FC; Exp, average expression; P-val, p-value.

5.7.3 Pagetoid SGC microRNA

Pagetoid expressed 53 significantly differentially expressed genes with a P value <0.05 (Figure 5.15). Just under half (24) are down regulated and all genes Log2FC ranged from -2.3 to 2.7. Expression ranged from 2.1 to 11 with only 15 genes expressing more than 5.0. The top 5 pagetoid specific genes are listed in Table 5.24.

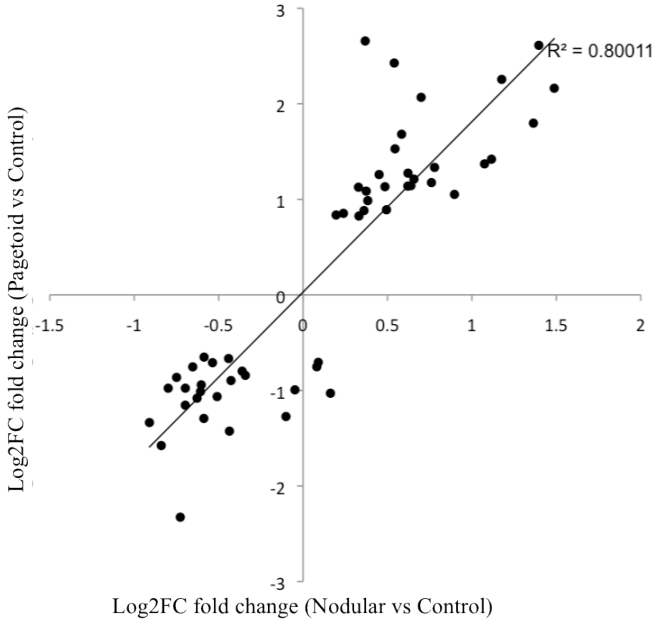


Fig 5.15 Pagetoid SGC specific microRNA. Fifty-three significant differentially expressed microRNAs unique to nodular SGC using a p <0.05 threshold.

	Pagetoid vs. Control			Nodular vs. Control			
miRNA	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P-val**	**Gene Target**
1308	2.7	5.9	0.014	0.4	5.9	0.695	
205	2.4	11.2	0.005	0.5	11.2	0.457	CDH1; ZEB1; ZEB2; *EZR; LMNA;* PTEN; PHLPP2; EZH2
200a	2.1	7.7	0.024	0.7	7.7	0.385	MAPK14; APP
1260	1.5	10.2	0.005	0.5	10.2	0.231	CCND1
199a-3p/199b-3p	-2.3	9.1	0.002	-0.7	9.1	0.217	CD44; AXL; NLK

Table 5.24 Pagetoid SGC specific microRNA. Top 5 differentially expressed genes present in pagetoid SGC only when normalised to tarsal plate control and using a p-value <0.05 and Log2FC fold change >1.5. Nodular expression shown to demonstrate non-significant change within the subtype. LC, Log2FC; Exp, average expression; P-val, p-value.

5.7.4 Comparison of nodular and pagetoid SGC microRNA

Identification of shared genes revealed 39 that were differential expressed in both subtypes compared to tarsal plate control (Figure 5.16). The majority are upregulated with only 12 downregulated and Log2FC ranged from -1.9 to 4.8. The top 5 DE genes are listed in table 5.25.

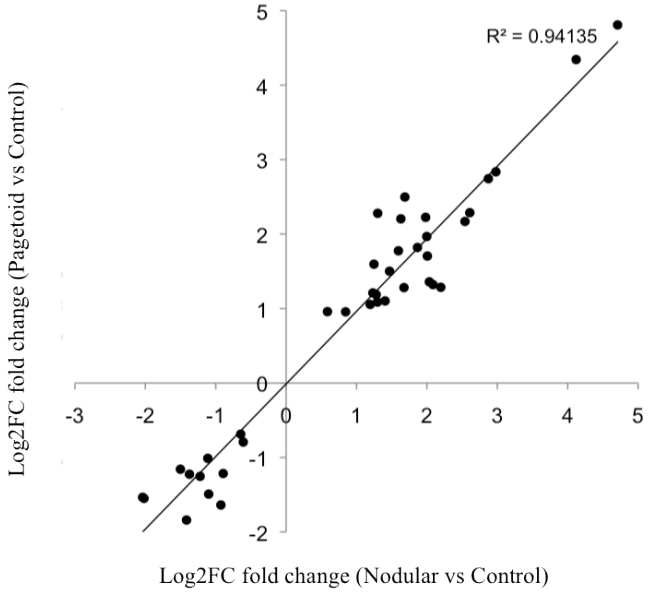


Fig 5.16 Shared microRNA between nodular and pagetoid SGC. Thirty-nine significant differentially expressed microRNA shared between both subtypes using a p <0.05 threshold.

	Nodular vs. Control			Pagetoid vs. Control			Gene Target
miRNA	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P-val**	
16	4.7	9.1	0.038	4.8	9.1	0.035	Bcl-2
9	4.1	6.2	<0.001	4.3	6.2	<0.001	CDH1
34a	3.0	5.4	<0.001	2.8	5.4	<0.001	Bcl2; Survivin; CDK4; CDK6; CCND1; CCNE1; MYC; E2F3; MET; SIRT1; MTA2; WNT1; LEF1; CD44
126	-1.1	10.9	0.021	-1.5	10.9	0.004	
619	-2.0	5.6	0.015	-1.5	5.6	0.047	

Table 5.25 Shared microRNA between nodular and pagetoid SGC. Top 5 differentially expressed genes present in both nodular and pagetoid SGC when normalised to tarsal plate control and using a p-value <0.05 and Log2FC fold change >1.5. LC, Log2FC; Exp, average expression; P-val, p-value.

5.7.5 Validation of RNAseq differentially expressed genes

Assessment of the driver mutations with SGC revealed no significant changes within the RNA array assessment. Only one gene came up in both techniques and that was SYCP2, which was 3.4/3.3 log2FC, 3.8/5.0 Exp, 8.1E-0.5/1.8E-0.5 P-value, for RNAseq/RNA array respectively. Two major families seemed affected though, the S100 and Histone proteins (Table 5.26). The latter, as mentioned, is in involved in chromosome structure and may reflect the chromosomal instability seen in the tumour.

	RNAseq			**RNA array**			**Subtype**
Gene	**LC**	**Exp**	**P-val**	**LC**	**Exp**	**P-val**	
SYCP2	3.4	3.8	8.1E-0.5	3.3	5.0	1.8E-0.5	Both
S100A14				1.4	7.3	0.002	Nodular
S100A8	3.2	9.3	0.03				
S100P	3.1	4.4	0.05				
HIST1H4B				2.0	4.3	0.008	Nodular
HIST2H2BE	0.6	4.3	0.223	1.9	5.6	0.0001	Nodular
HIST2H2AB				3.5	7.6	0.0002	Pagetoid
HIST2H2AB				1.9	6.1	0.002	Nodular
HIST1H2AC				1.0	4.9	0.06	Nodular
HIST1H2AC				1.7	5.6	0.004	Pagetoid
HIST1H2BI				1.7	4.6	0.0002	Pagetoid
HIST1H2AI				2.2	5.2	0.0001	Pagetoid
HIST1H3F				2.4	7.6	0.002	Pagetoid
HIST1H3I				3.0	6.2	0.0002	Pagetoid
HIST1H4L				1.8	3.8	0.0009	Pagetoid
HIST1H4L				1.1	3.2	0.02	Nodular
HIST2H4A				1.2	5.3	0.007	Nodular
HIST2H4A				1.3	5.5	0.003	Pagetoid
HIST1H4H	2.9	2.2	6.7E-0.3				

Table 5.26 Comparison RNAseq and Array expression. Expression seen in the histone and S100 family of genes demonstrating the burden across the 14 SGC samples, using a p-value <0.05 and Log2FC fold change >1.5. LC, Log2FC; Exp, average expression; P-val, p-value.

5.8 SGC MicroRNA validation using Taqman RT-qPCR

5.8.1 Endogenous eyelid miRNA control selection

For RNA RT-qPCR, there are well-documented housekeeping genes such as GAPDH or B-actin. The best method for deciding a housekeeping gene is to test the candidate in the tissues of interest to ensure similar expression is seen in all and that there is minimal fold change. For miRNA, there are no agreed universal housekeeping miRNA, so the expression levels of potential candidate housekeeping genes were assessed (Table 5.27) Hsa-miR-26a was highly expressed in both nodular and pagetoid SGC with minimal log2FC compared to normal eyelid tissue, so this was picked as the housekeeping gene.

	Nodular vs. control		**Pagetoid vs. control**	
MiRNA	**Log2FC**	**Expression**	**Log2FC**	**Expression**
miR-361-5p	-1.3	5.0	0.05	4.7
miR-186-5p	1.9	4.0	1.2	4.0
miR-26a-5p	-0.33	11	-0.12	11
miR-191-5p	1.3	6.3	2.3	6.3
miR-451a	-0.40	8.7	0.16	8.7
miR-423-5p	0.84	5.4	0.3	5.4
miR-320a	-0.6	3.1	-0.4	3.1

Table 5.27 Endogenous microRNA control selection. Listed are all possible endogenous controls for subsequent Taqman real time quantitative PCR with hsa-miR-26a-5p chosen as the control as it is highly expressed in both tissue subtypes and similarly activated.

5.8.2 Taqman miRNA qPCR

Verification of five differentially expressed miRNA from the Nanostring microarray was performed using Taqman RT-qPCR and hsa-miR-26a as the housekeeping gene (Figure 5.17). Hsa-miR-199a is significantly downregulated in both subtypes, but more so in pagetoid SGC which complements the array findings. Increased expression of hsa-miR-205 was seen in the pagetoid subtype, again supporting the array data. Both hsa-miR-150 and -34a are upregulated in nodular, but not pagetoid although there was a trend towards increased hsa-miR-34a expression. The former is a nodular specific miRNA on array scrutiny and this is validated on RT-qPCR. The later was supposed to be a shared miRNA. Hsa-miR-16 is highly expressed shared miRNA on both array and RT-qPCR testing.

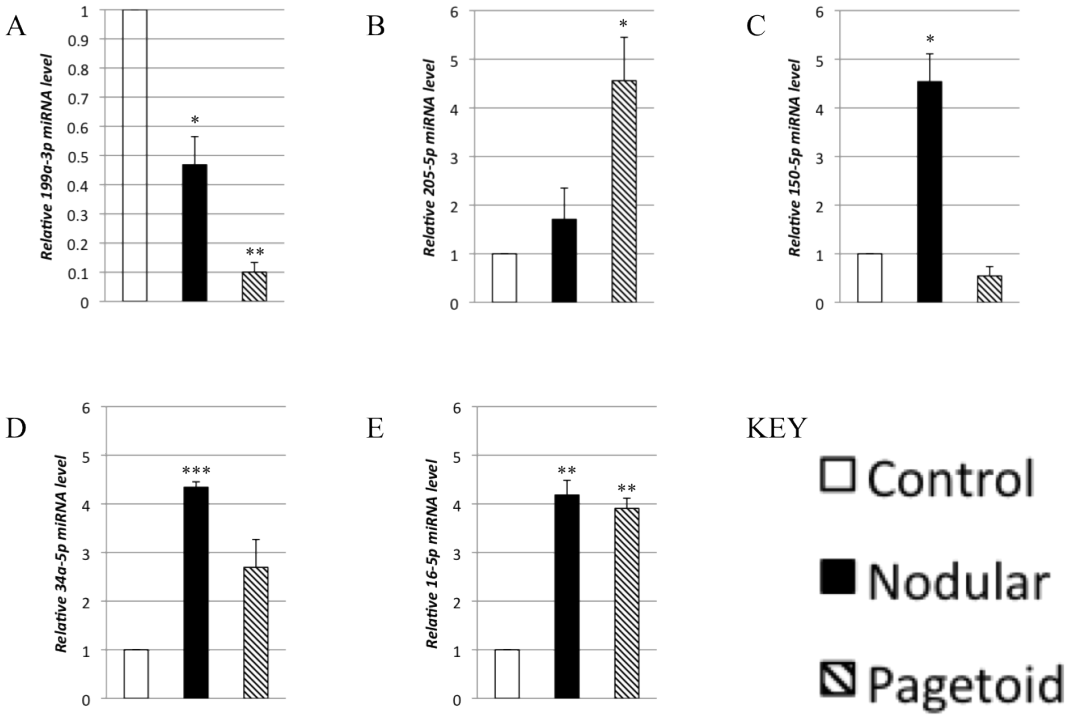


Fig 5.17 MicroRNA expression using Taqman RT-qPCR in SGC. Relative expression levels were determined for nodular and pagetoid SGC using Taqman RT-qPCR against normal eyelid tissue for miRNA A) 199a-3p, B) 205-5p, C) 150-5p, D) 34a-5p and E) 16-5p. *P<0.05, **P<0.02, ***P<0.01.

5.9 Protein Expression in SGC

5.9.1 Clinical features of SGC patients

Fifteen patients with SGC were included for comparison against 15 nodBCC (Hh pathway driven tumour control) patients with a mean age of 68 and 64, years respectively. All of the SGC patients had non-metastatic disease, 10 nodular and 5 pagetoid, requiring wide excision or exenteration as treatment.

5.9.2 Hedgehog expression in periocular SGC

Hh pathway expression was detected in all 15 SGC and nodBCC tumours as demonstrated by the DAB immunostaining for PTCH1, SMO, Gli1 and Gli2 (Figure 5.18). Western blot of LNCaP-Gli1 cells a known hyper-expressed Hh pathway cancer cell line, demonstrated good specificity of each antibody with a clear band for the appropriate sized protein (Fig 4.17, section 4.4.4). PTCH1 expression was detected in the cytoplasm of both nodBCC and SGC (Fig 5.18 E & I respectively) however, PTCH1 expression was markedly more pronounced in SGC compared to nodBCC. Similar levels of SMO were observed in both nodBCC and SGC (Fig 5.18 F & J). Gli1 and Gli2 were detected in the cytoplasm and nuclei of both nodBCC and SGC (Fig 5.18 G, K, H & I respectively). Gli1 expression was stronger in SGC compared to nodBCC whereas similar levels of Gli2 were observed in both nodBCC and SGC.

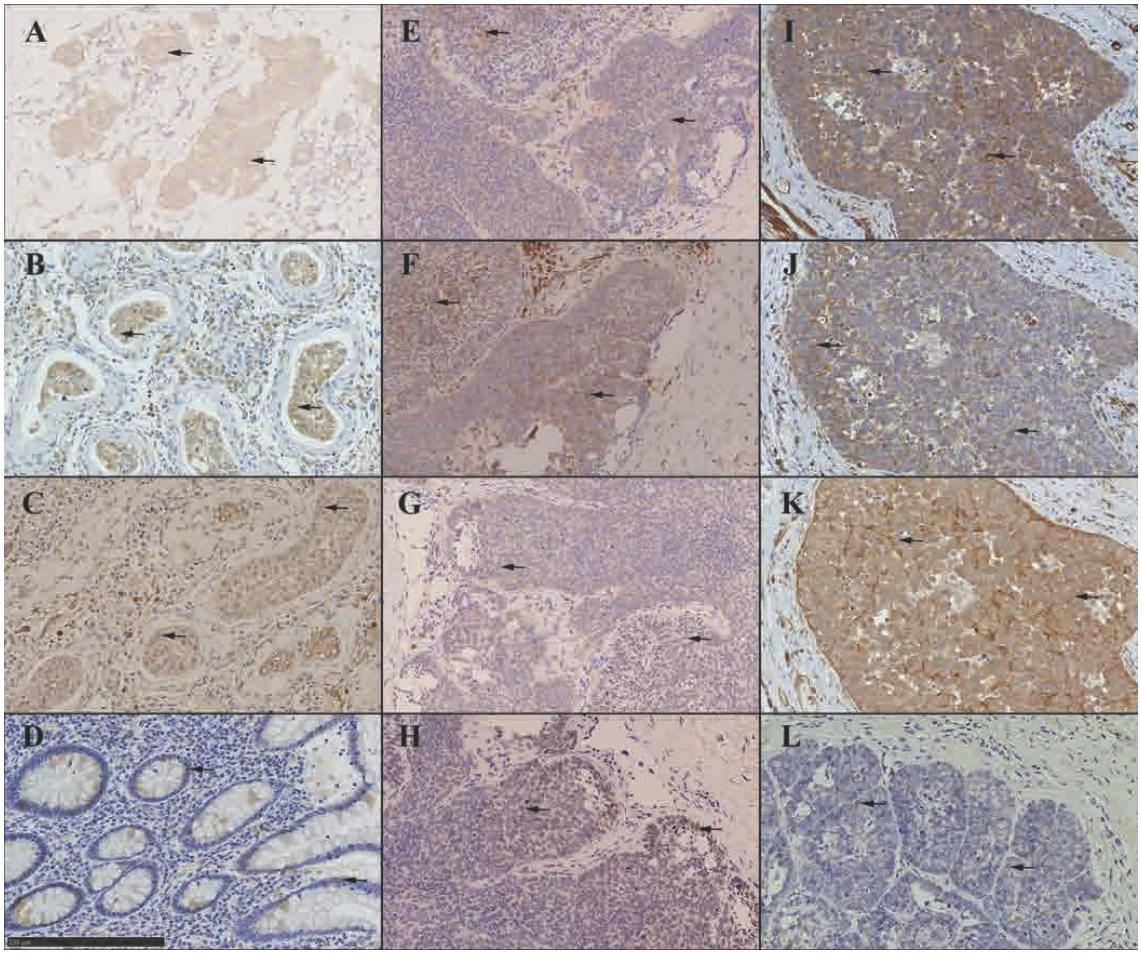


Figure 5.18 Hedgehog pathway expression in SGC using DAB immunostaining. Representative pictures of Hedgehog pathway expression in positive control tissue (A-D), nodBCC (E-H) and SGC (I-L). A, breast tissue displays PTCH1 expression in the cytoplasm of the cells (black arrows). B, SMO expression in testicular tissue, within the Leydig cells and seminiferous tubules cells (black arrows). C, Gli1 nuclear and cytoplasmic expression in testicular tissue, particularly in the Leydig cells and seminiferous tubules (black arrows). D, Nuclear expression of Gli2 in intestinal tissue, specifically localized to epithelial cells of the intestinal glands. E, PTCH1 expression in the cytoplasm (black arrows) of nodBCC, similar to control tissue. F, Similar SMO expression in nodBCC compared to control. G, Increased Gli1 nuclear expression in nodBCC compared to control tissue. H, Increased Gli2 nuclear expression compared to control tissue. I, Marked PTCH1 cytoplasmic expression in SGC compared to control and nodBCC. J, Slight increase in staining of SMO in the cytoplasm (black arrows) of SGC in contrast to nodBCC. K, Increased staining of nuclear (black arrows) and cytoplasmic Gli1 in SGC compared to nodBCC, and increased cytoplasmic expression compared to control. L, Nuclear staining of GLi2 (black arrows) in SGC that is similar to nodBCC, but more than control tissue. Scale bar represents 250 µm and all images are at 200 X magnification. Antibody stains for PTCH1 (A,E,I), SMO (B,F,J), Gli1 (C,G,K) and Gli2 (D,H,L). PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

DAB immunostaining is highly subjective as is the use of the immunoreactivity score and inter-observer variability has been shown in several studies. (Parker et al., 2002, von Wasielewski et al., 2002) Immunofluorescence was employed to semi-quantify the amount of protein expression in the SGC tumour in order to reduce this observer bias by computational counting along with the removal of background noise. Nevertheless, any immunohistochemistry method has many variables during processing so relative expression to another tumour processed in an identical fashion is a more robust way of analysing overall expression. In order to do this, a known aberrant Hh pathway tumour, nodBCC, was selected for comparison.(Saldanha et al., 2003) Hh signalling is normally switched off in the majority of adult tissues, however, some do transiently express Hh. These normal tissue samples represent physiologically expressed Hh and were also used for comparison to determine if expression was simply re-activation or aberrant overexpression.

Immunofluorescence analysis was used to better quantify the amount of PTCH1, SMO, Gli1 and Gli2 expression in SGC and nodBCC (Fig 5.19). A comparison between SGC tumour and nodBCC tumour was made along with physiologically activated Hh signalling (Fig 5.20). PTCH1, SMO, Gli1 and Gli2 displayed higher levels of fluorescence in SGC compared to nodBCC ($P<0.01$) (Fig 5.20A). PTCH1, SMO and Gli-2 were also expressed at higher levels in SGC than physiological Hh signalling ($P<0.01$). NodBCC had similar levels of PTCH1 and Gli2 to physiologically activated Hh, but markedly lower expression of SMO and Gli1 ($P<0.01$). Gli1 expression is lower in both tumours compared to normal expression; however, expression is still higher in SGC than nodBCC ($P<0.01$).

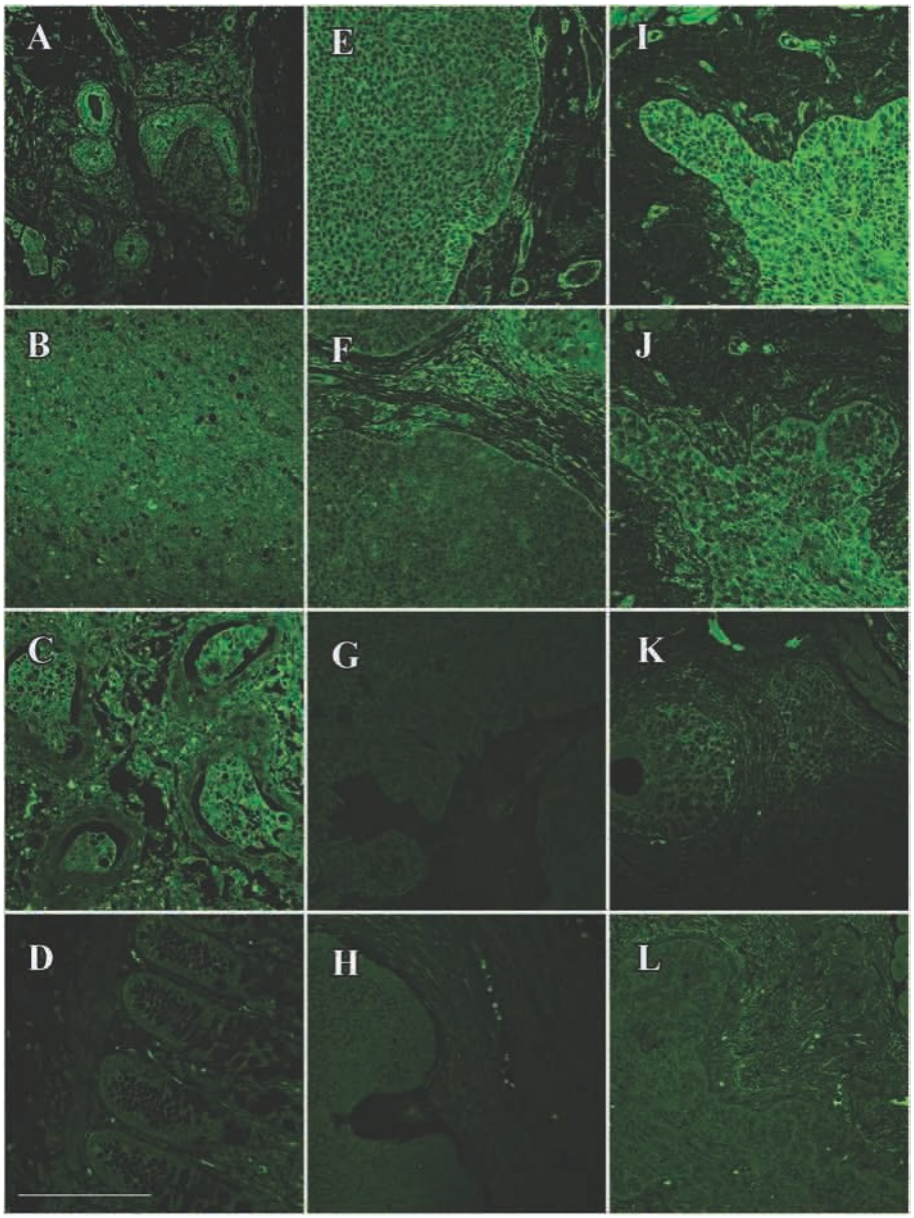


Figure 5.19 Hedgehog pathway expression in SGC using immunofluorescence. Representative pictures of Hedgehog pathway expression proteins using immunofluorescence in positive control tissue (A-D), nodBCC (E-H) and SGC (I-L). Secondary antibody staining utilized AlexaFluor-568 and the colour converted into black and white using Image J for pictorial purposes. Scale bar represents 250 μm and all images are at 200 X magnification. Antibody stains for PTCH1 (A,E,I), SMO (B,F,J), Gli1 (C,G,K) and Gli2 (D,H,L). PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

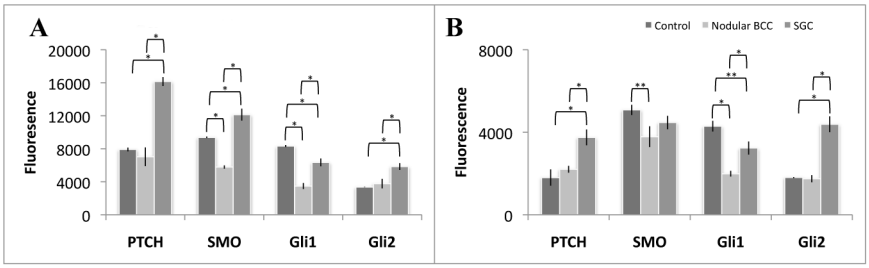


Figure 5.20. Hedgehog pathway semi-quantified expression in SGC using immunofluorescence. Semi-quantification of antibody expression (x-axis) within (A) SGC tumour, nodBCC tumour, control tissue and (B) stroma of SGC, nodBCC, control tissue using fluorescence intensity (y-axis) was determined in regions of interest as delineated by microscopy, ensuring a standardized area size whilst containing the same number of nuclei as determined by DAPI staining. Each tumour sample had an average of three readings. Each bar represents mean values ± SEM taken from 15 nodBCC, 15 SGC and control tissue samples for each Hh protein. *P<0.01, ** P<0.05. PTCH1=Patched 1; SMO=smoothened; Gli1=Glioma-associated zinc transcription factor1; Gli2= Glioma-associated zinc transcription factor2=Gli2.

Comparison of stromal expression using immunofluorescence (Fig 5.20B) showed that PTCH1, Gli1 and Gli2 were more highly expressed in the stroma of SGC compared to nodBCC (P<0.01). PTCH1 and Gli2 were highly upregulated in SGC stroma compared to both normal expression and nodBCC (P<0.01). Gli1 is less expressed in SGC (P<0.05) and nodBCC (P<0.01) compared to physiological Hh expression.

Levels of expression within SGC tumour were significantly higher than those observed in nodBCC. Furthermore, these levels were more than physiological re-activation except for Gli1 that was less than in normal Hh expressing tissue. It is estimated that up to 25% of tumours depend on Hh pathway activity for growth.(Lum and Beachy, 2004) Nodular BCC is a slow-growing cancer that metastasizes rarely.(Saldanha et al., 2003) In contrast, SGC is locally aggressive, metastasizes and has a risk of mortality of up to 29%.(Shields et al., 2004, Ni, 1982, Doxanas and Green, 1984) Thus, it is possible that the increased Hh expression seen in SGC accounts for its more aggressive nature compared to nodBCC. The exact role of Hh expression in SGC cannot be discerned from this study alone, as Hh signalling has been implicated in tumorigenesis, growth, progression, local invasion and metastasis.(Watkins et al., 2003, Thayer et al., 2003, Dahmane et al., 1997, Sanchez et al., 2005)

The stroma of SGC was found to express components of the Hh pathway and in particular, Gli2 was considerably more upregulated. Cancer can modify the surrounding non-malignant tissue to facilitate local invasion.(Marsh et al., 2008) Gli2 is an activator of the Hh pathway and plays a critical role in medulloblastoma tumorigenesis, with loss of Gli2 expression preventing the tumour formation.(Hui and Angers, 2011, Flora et al., 2009) Moreover, higher levels of Gli2 are associated with a loss of the cell adhesion molecule E-cadherin in melanoma cell lines with increased capacity for local invasion, higher levels correlated with the most aggressive tumours. (Alexaki et al., 2010) In contrast, knockdown models of Gli2 demonstrated reduced tumour growth in prostatic in vivo models.(Thiyagarajan et al., 2007) Hh expression as a lone stimulator is unlikely and crosstalk with multiple pathways can occur, for example, transforming growth factor-β (TGF-β) has been shown to promote Gli2 expression and is involved in tumour progression and metastasis.(Javelaud et al., 2011)

5.9.3 HIST1H2 expression in periocular SGC

Expression of HIST1H2BM within SGC was confirmed at a protein level with strong levels seen within all samples tested (Fig 5.21)

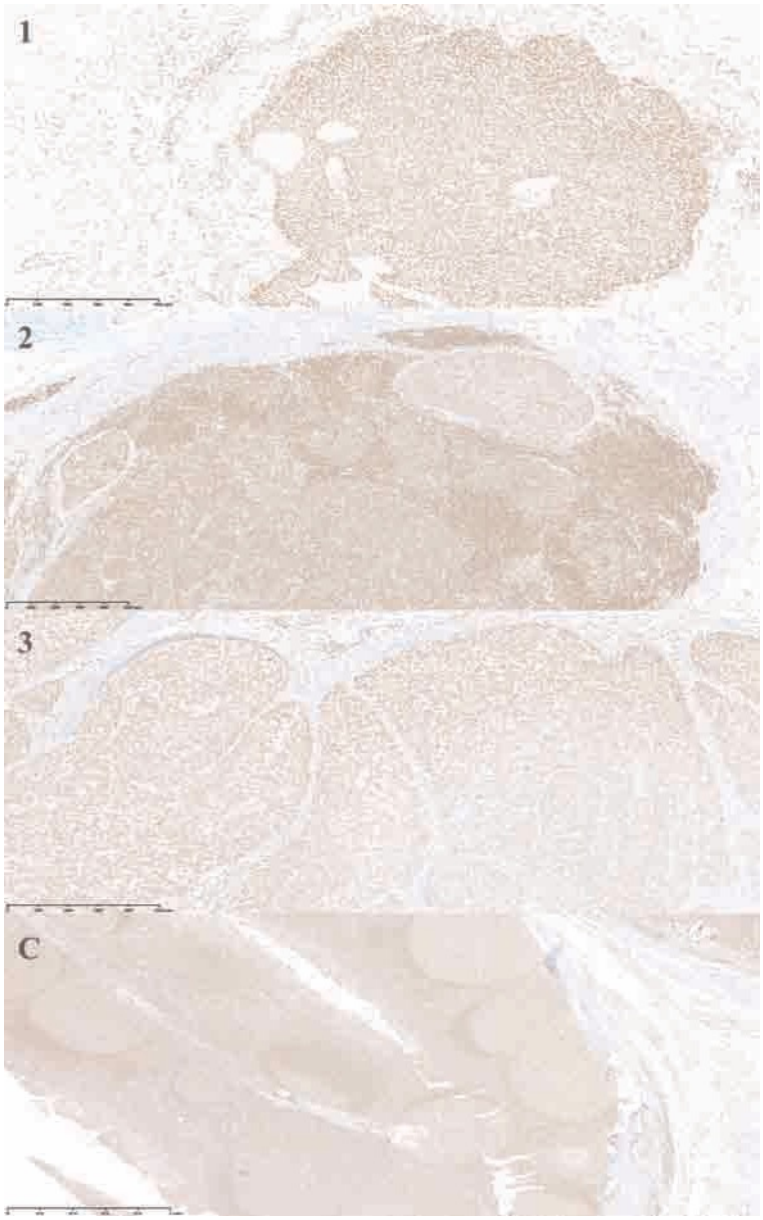


Fig 5.21 HIST1H2BM expression in SGC using DAB immunostaining. Three SGC samples were analyses and compared to control (C, control tonsillar tissue).

DISCUSSION

5.10 Whole exome sequencing in SGC

5.10.1 Shared driver mutations using MutSigCV and Intogen

Retinoblastoma 1 (*RB1*) – chromosome 13q14.1

RB1 is also an archetypal tumour suppressor gene that causes the hereditary childhood cancer, retinoblastoma (RB). The RB1 protein binds to the E2F family of transcription factors and represses the transcription of genes needed for the S phase of the cell cycle, thus, is a repressor of the cell cycle. (Nevins, 2001) Furthermore, RB is a paradigm for the Knudson's two hit hypothesis whereby loss of both *RB1* genes are required for the retinoblastoma to develop.(Cavenee et al., 1983) Nevertheless, loss of RB1 on both alleles although necessary, is likely not to be sufficient alone to initiate tumorigenesis. How loss of RB1 initiates cancer in retinoblastoma is probably dependent on the cell of origin assuming that cancer is a clonal expansion of a single cell in the first instance. Its role is more diverse that a tumour suppressor gene and may play a direct role in mitosis, genome stability via chromatin regulators, interact with ubiquitin ligases and more recently, cellular senescence.(Chinnam and Goodrich, 2011) The latter, results in permanent or long term prevention of cell proliferation in response to stress, provides a barricade to tumour initiation and both RB1 and TP53 are key TSGs required for enforcing this senescence barricade.(Courtois-Cox et al., 2008) Despite the variety of cancer inducing senescence mechanisms (such as DNA damage, oxidative stress, reactive oxygen species, heterochromatin formation) with the associated intermediate pathways involved, all routes pass through RB1 and TP53 control. Loss of both pathways in SGC may account for its highly aggressive nature. Small cell lung cancer is a highly aggressive cancer with a 5% 5 year survival and only contains few mutations, but TP53 plus RB1 are recurrently mutated.(Fiorentino et al., 2016) Moreover, functional inactivation of both RB1 and TP53 can initiate small cell lung cancer in mice.(McFadden et al., 2014)

Replication initiator 1 *(REPIN1)* – chromosome 7q36.1

REPIN1 was first described as a DNA binding protein involved in DNA bending, but more recently shown to regulate adipocyte size and expression of glucose transporters.(Ruschke et al., 2010) High levels of REPIN1 (AP4) have been correlated

with a poor prognosis in colorectal cancer, breast and prostate cancer.(Jung et al., 2008b) Specifically, it upregulates the genes involved in EMT and cell proliferation.(Chen and Chiu, 2015) Moreover, it regulates MMP9, a metalloproteinase involved in stromal breakdown, thus promotes cell migration, which is an enhanced characteristic of SGC.(Ku et al., 2009) Interestingly, mutant TP53 and REPIN1 have shown to promote cell movement and invasion.(Chen and Chiu, 2015)

Zinc finger protein 37A *(ZNF37A)* – chromosome 10p11.1

Little is known about this zinc finger protein other that it is likely to be involved in DNA binding and transcription activity. However, the region that it sits in (the central region of chromosome 10) is also where genes for the inherited cancer syndrome reside: multiple endocrine neoplasia or MEN.(Tunnacliffe et al., 1993)

Fragile Site, Folic Acid Type, Rare, Fra(10)(Q23.3) Or Fra(10)(Q24.2) Candidate 1 *(FRA10AC1)* – chromosome 10q23.33

Minimal data exists on this gene other that its association with amyloid in the cerebrospinal fluid with possible links to Alzheimer's disease.(Yao et al., 2015)

5.10.2 Drivers identified using MutSigCV algorithm

Integrin Binding Sialoprotein (*IBSP*) – chromosome q22.1

An elevated level has been seen in oesophageal squamous cell carcinoma and is thought to play a role in malignant progression with its presence related to a poorer 5-year survival compared to those not expressing IBSP.(Tang et al., 2014) Higher levels confer a poorer prognosis in renal cell carcinoma especially when co-expressed with osteopontin.(Righi et al., 2013) These comprise of a group entitled small integrin binding ligand N-linked glycoproteins (SIBLINGS) who mediate bone morphogenesis and may facilitate the metastasis of primary cancers to bone.(Kruger et al., 2014) SGC tends to metastasis to the lymph nodes first, but has been shown to go to the lung, liver and bone.(Husain et al., 2008)

Chromosome 16 Open Reading Frame 87 (*C16orf87*) – chromosome 16q11.2

Nothing is known about this open reading frame.

Interleukin 12 Receptor Subunit Beta 2 (*IL12RB2*) – chromosome 1p 31.3

The IL-12 receptor consists of two subunits: IL12RB1 and 2. Cytokine IL-12 has been shown to regulate the inflammatory process and be part of the host immune defence. It has anti-tumour activity and can suppress UV radiation induced apoptosis.(Isomura et al., 2008) Lack of signalling via IL12RB2 predisposes to autoimmunity and malignancy with loss of function found in B cell malignancies.(Airoldi et al., 2005) Polymorphisms within this gene confer risk of developing lung adenocarcinoma.(Prigione et al., 2016)

Glutathione Peroxidase 3 (*GPX3*) – chromosome 5q33.1

GPX3 plays an important role in the detoxification of hydrogen peroxide and is the only extracellular member of this family highlighting its importance in extracellular defence against the effects of reactive oxygen species. In vitro GPX3 was shown to reduce colitis-induced carcinoma.(Barrett et al., 2013) Loss of function via hypermethylation has been demonstrated in several cancers including cervical, breast, gastric, oesophageal SCC and leukaemia.(Yao et al., 2015) Anti-invasiveness and growth effects of GPX3 have been shown to be mediated by deactivation of Erk-NFκB-SIP1 signalling in hepatocellular carcinoma confirming its role as a tumour suppressor gene.(Qi et al., 2014)

Repetin (*RPTN*) – chromosome 1q21.3

Epidermal differentiating genes have been clustered around chromosome 1q21.3 and RPTN is a member of the fused gene family that clusters around this region. Fused members including profilaggrin, trichohyalin and hornerin are associated with keratin intermediate filaments acting as a barrier in normal skin.(Huber et al., 2005) In addition, it is a member of the S100 family and can reversible bind calcium. Increased expression of repetin is a biomarker for susceptibility to inflammation.(Xu et al., 2014) Repetin is not just localised to the epidermis and expression has been shown in the brain with diminished levels associated with schizophrenia and bipolar disorder.(Wang et al., 2015c)

Gap Junction Protein Delta 3 (*GJD3*) – chromosome 17q21.2

This is a member of the connexin family involved in GAP junction formation. These intercellular communications are often lost in cancer and they act as tumour suppressor genes. Loss of function via hypermethylation of connexin including *GJD3* has been shown in colorectal cancer.(Sirnes et al., 2011)

Distal-Less Homeobox 6 (*DLX6*) – chromosome 7q21.3

DLX6 has an important role during embryogenesis and works in unison with DLX5 in pharyngeal arch formation, especially establishing proximal-distal identity.(Depew et al., 2002) They are also involved in endometrial maturation post-natally.(Bellessort et al., 2016) Although DLX6 has not been reported in caners, the DLX family does play a role in promoting tumour growth and progression.(Li et al., 2015b)

G Protein Subunit Gamma 2 (*GNG2*) – chromosome 14q22.1

The heterotrimeric G protein comprises of a GNG2 dimer and G alpha subunit. This G protein is involved in a angiogenesis, differentiation, invasion and proliferation.(Schwindinger and Robishaw, 2001) Overexpression was shown to inhibit melanoma cell metastasis via a decrease in focal adhesion kinase activity.(Yajima et al., 2014)

G Protein-Coupled Receptor 31 (*GPR31*) – chromosome 6q27

This G protein receptor has a high affinity for 12-(S)-hydroxy-5,8,10,14-eicosatetraenoic acid (12-S-HETE), an arachidonic acid metabolite secreted by platelet and tumour cells.(Guo et al., 2011) Increased expression of 12-S-HETE positively correlates with aggressive prostate cancer, as does GPR31 expression.(Honn et al., 2016) GPR31 expression is an independent risk factor for poorer prognosis in colorectal cancer.(Zou et al., 2015)

5.10.3 Drivers identified using Intogen algorithm

Titin (*TTN*)– chromosome 2q31.2

See section 3.2.3.iii. As mentioned, this is unlikely to be of biological significance.

Pre-MRNA Processing Factor 31 (*PRPF31*) – chromosome 19q13.42

PRPF31 is a splice site factor in pre-mRNA processing and required for tri-SNP formation along with its spliceosome activity.(Makarova et al., 2002) Mutations within this gene results in autosomal dominant retinitis pigmentosa (RP) and highlights that pre-mRNA splicing is an important route to disease.(Liu and Zack, 2013) Polymorphisms within PRPF31 are associated with a higher risk of invasive breast cancer.(Peedicayil et al., 2010) Rescue of this gene has been shown using non-viral gene therapy

(nanoparticles) in the murine model and this has potential for SGC patients in the future if a loss of function is tumorigenic.(Pensado et al., 2016)

Cubilin (*CUBN*) – chromosome 10p13

This codes for a gene that is frequently located at absorptive epithelial layers and is a co-transporter for vitamin, iron and lipoprotein.(Aminoff et al., 1999) Mutations in CUBN results in albuminuria.(O'Toole and Sedor, 2011) GWAS have shown genetic susceptibility in CUBN to gastric cancer and colorectal cancer.(Zhao et al., 2016, Al-Tassan et al., 2015)

SYNE1 – Chromosome 6q25.2

See section 3.2.3.ii. SYNE1 is a nuclear membrane protein involved in structural cellular organisation in muscle and the central nervous system. It may not be biologically relevant; however, methylation has been shown in colorectal cancer and SNVs are associated with invasive ovarian cancer.

5.11 RNA sequencing in periocular SGC

5.11.1 Upregulated genes in SGC

Small Proline Rich Protein 2A (*SPRR2A*) – Chromosome 1q21.3

This was initially described as a keratinocyte protein involved in structural integrity but is also a stress and wound healing modulator by reducing reactive oxygen species and allowing transient EMT.(Mizuguchi et al., 2012) In liver disease, including cancer, it promotes EMT of biliary epithelium via ZEB-1 as part of normal and pathological wound healing response.(Mizuguchi et al., 2014) SPRR2A expression in cholangiocarcinoma induces permanent EMT, aids tumour growth and creates a permissive environment for local invasive behaviour – a key feature of pagetoid SGC.(Specht et al., 2013) In addition, SPRR2A is a repressor of P53 dependent transcriptional activity by interfering with TP53-EP300 activity and increasing HDAC1 expression.(Mizuguchi et al., 2012) This region was amplified on chromosomal analysis.

Small Proline Rich Protein 2D (*SPRR2D*) – chromosome 1q21.3

Similar to the above-related family member, this protein is expressed in both neoplastic and inflammatory disease.(De Heller-Milev et al., 2000)

Serpin Family B Member 4 (SERPINB4)– Chromosome 18q21.33

Serine protease inhibitor family B inhibit granzymes released by NK cells and cytotoxic T lymphocytes that induce cell death as part of the innate immune system.(Smyth et al., 2001) One such granzyme, GrM, is a potent inducer of tumour cell death that SERPINB4 inhibits by forming a stable complex; expression of SERPINB4 by tumour cells shields them from GrM apoptosis along with NK cell death.(de Koning et al., 2011) Oropharyngeal squamous cell carcinoma and primitive neuroectodermal tumours of the brain both express SERPINB4.(van Kempen et al., 2016, Vermeulen et al., 2016)

Serpin Family B Member 3 (*SERPINB3*) – Chromosome 18q21.33

SERPINB3 is another member of the family B serpins. It was originally described as a squamous cell antigen in those who had cervical SCC, and subsequently it is associated with lung, liver plus head and neck cancer.(Turato and Pontisso, 2015) It augments liver carcinogenesis by increasing transcription of MYC through the YAP pathway.(Turato et al., 2015) It is associated with both the TGF-B and Wnt signalling to create a tumorigenic microenvironment that confers a poor prognosis in liver cancer.(Pontisso et al., 2014)

Alanine And Arginine Rich Domain Containing Protein (AARD)– Chromosome 8q24.11

Little is known about this gene other than it has a predilection for testicular tissue.(Svingen et al., 2007)

Synaptonemal complex protein 2 (*SYCP2*) – chromosome 20q13.33

Synaptonemal protein complexes are required for the joining of homologous chromosomes during the meiotic prophase stage. Specifically, during zygotene (the second stage of prophase), homologous chromosomes come together allowing inter-homolog repair of the preceding stage's programmed double stranded breaks (DSB). A feature of this repair mechanism is the tripartite protein structure formed by the synaptonemal complex (SC) that synapses homologous chromosomes along their length. *SYCP2* works in concert with the lateral element protein SYCP3 and the cohesion protein RAD21. Furthermore, SYCP2 acts as the linker between SYCP3 and SYCP1, hence binds

the lateral elements and the transverse filaments; an essential step for proper chromosome synapses. (Winkel et al., 2009) Another possible role for SC complexes is centromere coupling. Overexpression occurs in HPV positive oropharyngeal squamous cell carcinoma.(Martinez et al., 2007) Interestingly, this was also detected in the separate SGC samples below using array analysis, thus validating this as an important differentially expressed gene in SGC.

S100 Calcium Binding Protein A8 (*S100A8*) – chromosome 1q21.3

S100A8 is a zinc and calcium binding protein involved in the regulation of inflammatory and immune response. It is part of the S100 protein family that regulated both intracellular and extracellular homeostasis mechanisms and when released into the extracellular matrix can promote sustained inflammatory damage.(Lim et al., 2016) Specifically, it promotes migration of colorectal cancer cells via the NF-KB pathway in macrophages.(Zha et al., 2016) Moreover, S100A8 is essential for cell migration and invasion but not necessary for proliferation.(Lim et al., 2016)

S100 Calcium Binding Protein P (*S100P*) – chromosome 4p16.1

S100P has a similar general function as mentioned above with the P standing for placenta as the original tissue of detection. S100P however, is more associated with metastatic disease and plasma levels are related to a higher chance of metastasis in breast cancer, but these levels come down with subsequent treatment making S100P a potential metastatic cancer marker.(Peng et al., 2016) Similar to S100A8, it also promotes cell proliferation via the NF-KB pathway with an extracellular interaction with receptor for activated glycation end products (RAGE).(Prica et al., 2016) It has a multitude of roles in carcinogenesis and therapeutic blocking agents have been used including cromolyn, an anti-allergic drug.(Prica et al., 2016, Arumugam et al., 2013)

Histone Cluster 1, H4h (*HIST1H4H*) – chromosome 6p22.2

Histones are part of the wrapping of DNA so it can be packed into a small space. An octamer of histones (two H2A-H2B dimers with a H3-H4 tetramer) has 146 bp of DNA wrapped around it (termed the nucleosome) and linked to other nucleosomes with linker DNA.(Grunstein, 1990) Linker histones sit at the base of the nucleosome and result in a larger increase in DNA supercoiling as well as chromatin condensation.(Khochbin, 2001) The nucleosomes look like beads on string (euchromatin) and this is further wrapped

using linker histones (such as histone H1) to form a supercoiled structure heterochromatin (during gene inactivity).(Smith and Santisteban, 1998) The exact structure of chromatin (ultimately a chromosome) depends on the exact phase of the cell cycle: the metaphase chromatin is known as the classic chromosome (duplicated chromosome) seen during mitosis when they became visible on microscopy (karyotyping). Histone modification is a potential therapeutic target and targeting of the methylation sites on histones occurs with Temozolomide, a DNA methylating chemotherapy agent and could potentially be used in SGC patients.(Wang et al., 2016a)

Epithelial Cell Adhesion Molecule (EPCAM) – chromosome 2p21

EPCAM is a transmembrane cell adhesion protein and expressed in the majority of epithelial tissue. Consequently, overexpression is seen in many epithelial caners including colorectal, breast, gastric, liver and prostate.(Andree et al., 2016) Higher levels of expression are correlated with a poorer prognosis in breast cancer.(Ohashi et al., 2016) Due to its widespread prevalence in epithelial cancers, antibodies have been developed to target EPCAM and phase 1 trials have occurred using Catumaxomab in peritoneal metastasis, however, significant side effects do occur including fatal hepatitis.(Borlak et al., 2016)

Fatty Acid Binding Protein 5 (*FABP5*)– Chromosome 8q21.3

FABP5, whose role is in fatty acid transport, is expressed in all tissues of the body despite first being detected in keratinocytes and overexpressed in psoriasis.(Ogawa et al., 2011) Overexpression is seen in cancers and confers a poorer prognosis in triple negative breast cancer.(Liu et al., 2011a) Expression of FABP5 in concert with specificity protein 1 (SP1) and MYC is important in prostate carcinogenesis.(Kawaguchi et al., 2016) Together FABP5 and peroxisome proliferator-activated receptor delta (PPAR∂) enhance proliferation, migration and invasion in breast cancer.(Levi et al., 2013)

5.11.2 Downregulated genes in SGC

Phosphorylase, Glycogen, Muscle (PYGM) – chromosome 11q13.1

PYGM is an enzyme that resides in muscle to carry out glycogenolysis. Mutations in PYGM results in a glycogen storage disease, McArdle disease or myophosphorylase deficiency.(Nogales-Gadea et al., 2015) WES of rare aggressive breast cancer revealed

PYGM mutations in 30% of cases and is downregulated in several cancers suggesting a role for glycogen metabolism in cancer evolution.(Dieci et al., 2016)

Chromogranin B (CHGB) chromosome 20p12.3

The neuroendocrine system contains secretory granules which encompass chromogranin A and B. CHGA has been shown to be useful to monitor neuroendocrine tumours, whereas CHGB is less important.(Monaghan et al., 2016) Low CHGB levels in pancreatic neuroendocrine tumours have been associated with malignant tumours and metastasis.(Weisbrod et al., 2013) The relevance to SGC is unclear.

Immunoglobulin Heavy Constant Gamma 1 (G1m Marker) (IGHG1) – chromosome14q32.33

Immunoglobulin is a Y shaped protein secreted by plasma cells (antibody) or found on the surface of cells. The two ends of Y contain an area that binds to a specific antigen (the paratope) and is hypervariable in order to bind to lots of different antigens. Four polypeptide chains make up the Y region, two heavy and two light chains. The heavy chain consists of a variable and constant domain, as does the light chain and *IGHG1* codes for the constant heavy chain. IGHG1 is the most important isoforms of the immunoglobulins produced by the body and binds efficiently to effecter cells such as macrophages, NK cells, neutrophils and triggers the humeral immune response that kills cancer cells. Loss of expression would seem to confer a benefit for SGC to evade the body's immune system. Nevertheless, several studies have shown an increase in IGHH1 within pancreatic cells aids proliferation and silencing IGHH1 in prostate cancer cells promotes apoptosis, although the mechanisms for benefiting the cancer are unclear.(Li et al., 2011)

CD74 Molecule (CD74) – chromosome 5q33.1

CD74 plays an essential role in major histocompatibility class 2 antigen presentation complex. Expression has been detected in many cancers with higher expression levels in gastric cancer associated with a poorer prognosis.(Ishigami et al., 2001) Blocking of CD74 has been attempted in B-cell malignancies with the development of Milatuzumab which has passed phase 1 trials in 2010; however, nothing has been heard since these trials.(Berkova et al., 2010) It highlights that caution should be taken as to block something just because it is over-expressed without knowing its biological significance

could be risky. SGC is one of the most aggressive tumours, so that loss of CD74 with immune evasion may be an important factor.

Gelsolin (GSN) – chromosome 9q33.2

Gelsolin is required for cytoskeletal turnover and regulates protease secretion. Specifically, it acts to breakdown actin in concert with vimentin and together are normally up regulated to promote tumour growth and invasiveness.(Huang et al., 2016a) Nevertheless, GSN can be both an inhibitor and effector of apoptosis; GSN can enhance DNAse1 and apoptotic activity through the GSN-HIF1-a-DNase 1 pathway.(Li et al., 2012a) Decreased expression has also been seen in a variety of cancers including breast, prostate, gastric, kidney and so it is thought that it may act as a tumour suppressor in certain situations.(Winston et al., 2001) Lower levels of GSN in bladder cancer conferred a poorer prognosis and predominantly lower levels were seen in the presence of TP53 mutations.(Sanchez-Carbayo et al., 2007)

Collagen Type I Alpha 2 (*COL1A2*) – chromosome 7q21.3

Collagen Type I Alpha 1 (*COL1A1*)– chromosome 17q21.33

Type I collagen contains two alpha 1 chains and one alpha 2 chain to form a triple helix, *COL1A1* codes for alpha 1 and *COL1A2* codes for alpha2. During cancer transformation in vitro, collagen production is decreased. The more indolent solid tumours tend to have a surrounding stroma comprising of type 1 collagen.(Smith et al., 1983) Loss of surrounding type I collagen may confer a benefit for locally invasive tumours and *COL1A2* is frequently silenced by hypermethylation in cancer.(Sengupta et al., 2003) This loss of a collagen capsule may allow for faster SGC growth and induce the pagetoid behaviour that is seen clinically allowing the tumour to seed indiscriminately. A decrease in COL1A1 is seen in liver cancer with a return to expression after chemotherapy suggesting the potential for COL1A1 to be a biomarker with its presence indicating remission.(Hayashi et al., 2014)

Collagen Type III Alpha 1 (*COL3A1*)– chromosome 2q32.2

Type III collagen is frequently associated with type I collagen within connective tissue so it is not surprising that both are down regulated together. A study into nasopharyngeal carcinoma showed the enhancement of invasive behaviour with hsa-miR-29a/b, which in turn regulates COL3A1.(Qiu et al., 2015)

Apolipoprotein E (APOE)– chromosome 19q13.32

APOE plays a role in lipid metabolism and protein synthesis, but is also involved in tissue repair, immune response and cell growth. Overexpression has been shown in gastric cancer.(Shi et al., 2015a) Genetic polymorphisms within the gene have demonstrated an increase risk in renal cell carcinoma, but not testicular cancer.(Liu et al., 2015a, Lv et al., 2015)

Decorin (DCN) – chromosome 12q21.33

DCN is a proteoglycan who plays a role in stimulating autophagy, inflammation and inhibits tumorigenesis, so acts as tumour suppressor. It is often secreted by fibroblasts and modifies cellular behaviour. Loss of DCN allows renal cell carcinoma proliferation and metastasis via downregulation of P21 and E-cadherin.(Xu et al., 2016) Gain of function KRAS mutations have been shown to suppress DCN to aid colorectal cancer growth.(Mlakar et al., 2009) DCN mediates its anti-cancer effects by antagonising TGFB.(Qian et al., 2014) Loss of this TSG in SGC may confer a benefit.

Muscle related genes of unknown significance

Myosin, Heavy Chain 4, Skeletal Muscle (MYH4) 17p13.1

This is a skeletal muscle coding protein and biologically is unlikely to be related cancer. It is however, located at the same chromosomal location as TP53 which is preferentially lost in SGC and hence is probably a passenger.

Parvalbumin (PVALB) – chromosome 22q12.3

PVALB is involved in muscle relaxation, however, may be a useful tumour marker to detect conversion from benign adenoma to carcinoma in thyroid cancer.(Cerutti et al., 2011)

ATPase Sarcoplasmic/Endoplasmic Reticulum Ca2+ Transporting 1 (ATP2A1) – chromosome 16p11.2

This is an enzyme involved in muscle excitation and unlikely to be of significance.

Myosin, Heavy Chain 1, Skeletal Muscle, Adult (MYH1) – chromosome 17p13.1

MYH1 is located near TP53 and this region is preferentially lost in SGC.

Sarcalumenin (SRL) – chromosome 16p13.3

SRL is a muscle related gene and unlikely to be related to cancer growth.

Creatine Kinase, M-Type (*CKM*) – chromosome 19q13.32

CKM muscle specific creatine kinase.

Calsequestrin 1 (*CASQ1*)- chromosome 1q21

This is a skeletal muscle protein.

Tropomyosin 2 (Beta) (*TPM2*) - chromosome 9p13.3

TPM2 is a gene involved in slow muscle fibre contraction and mutations can result in myopathy.

Titin (*TTN*)– chromosome 2q31.2

See section 3.2.3.iii

5.12 SGC RNA subtype analysis

5.12.1 Nodular SGC messenger RNA

RNA, U7 small nuclear 141 pseudogene (RNU7-141P) - chromosome 20q13.33

A pseudogene is a sequence that is related to a known gene but does not have the ability to be expressed or code for a polypeptide, hence falls under the family of a non-coding RNAs. Small non-coding RNAs include the well-known transfer RNA (tRNA) and uracil rich small nuclear RNAs such as small nuclear, small nucleolar, transfer RNA (tRNA) and Y RNAs (see below). The U symbol is used to denote their nuclear location.(Rogers, 1985) In general, the formation of processed pseudogenes requires the expression of both L1 proteins.(Esnault et al., 2000) They are classed as a dysfunctional group, but can retain promoter or splice site activity, however, their biological activity is not clear. With over 40% homology with functional genes, they are described as ancestral remnants of a gene suggesting they diverged as part of an evolutionary event.(Wei et al., 2001) In some cases, their non-functionality is derived from the fact they contain multiple mutations such as premature stop codons or frameshifts, thus may be an evolutionary by-product of

accumulated mutations. Although pseudogenes are inactive by definition, they can still be inserted into the genome and unwittingly contribute to exons of other genes through alternative splicing of a transcript. Rarely, they undergo gene resurrection by undergoing gene conversion, whereby there is a mismatch during homologous recombination. In the same manner, the pseudogene can act a template for a variant for a gene that could be advantageous as well as deleterious.

Y RNA *(Y_RNA.52-201)* – chromosome 6q15

This is a small non-coding 96 bp RNA sequence. Initially Y RNAs were described as being in the cytoplasm of the cell, hence given the Y prefix in contrast to the aforementioned U RNA which were discovered in the nucleUs when looking at the associated antigens Ro60 and La of systemic lupus erythematosus.(Lerner et al., 1981) The Y RNAs have binding sites for proteins and the function of this complex is implicated in RNA processing or quality control mechanisms. (Wolin et al., 2012) Although Y RNAs are part of the small non-coding RNA group too, they have a known essential biological role in the initiation of chromosomal DNA replication. (Krude et al., 2009) Blocking of Y RNAs in *Xenopus laevis* (African claw frog) and *Danio rerio* (zebrafish) embryos is lethal due to abrogation of DNA replication, arresting of the cell cycle followed by cell death.(Collart et al., 2011)

Synaptonemal complex protein 2 (*SYCP2*) – chromosome 20q13.33

See section above 5.4.2 Detection of SYCP2 using a second modality validates this as an important gene in SGC tumourigenesis.

Small nucleolar RNA, C/D box 51 (SNORD51) – chromosome 2q33

Small nucleolar RNA (snoRNA) are non-coding RNAs implicated in the modification of tRNA, ribosomal RNA and small nuclear RNAs. C/D box snoRNAs are involved in methylation, hence can control gene expression. SNORD50A and 50 B (both co-localised on Chromosome 16) are recurrently deleted in human cancer and their loss is associated with reduced survival. (Siprashvili et al., 2016) Specifically, SNORD50 suppresses the activity of RAS oncoproteins and loss of them in cancer enhances KRAS activity along with ERK1/2 MAPK pathway.(Siprashvili et al., 2016)

Small Cajal body-specific RNA 8 (*SCARNA8*) – chromosome 9p22.1

SCARNA8 is a type of snoRNA localised to sub-organelles (cajal bodies) within the nucleus of proliferative cells, thus may reflect the presence of excessive cell division.

RNA, U7 small nuclear 57 pseudogene (*RNU7-57P*) – chromosome 1q21.3

This is a pseudogene of unknown significance.

Small Cajal body-specific RNA 1 *(SCARNA1)* – chromosome 1p35.3

This is a snoRNA of unknown significance

5.12.2 Pagetoid SGC messenger RNA

Immunoglobulin kappa joining 3 (*IGKJ3*) - chromosome 2p11.2

The light chain consists of a variable and constant domain (as do the heavy chain) and can be one of two types, either lambda or kappa chain, however, both cannot be present on the same antibody (i.e. two lambda or two kappa). In order to generate a unique immunoglobulin, the variable region is encoded by several subgenes that undergo somatic recombination. These subgenes include variable (V), diversity (D) and joining (J) segments (such as *IGKJ3)* that are tandemly arranged across the DNA. The B-lymphocyte will randomly select these subgenes to arrange a unique variable region to a specific antigen (known as V(D)J or somatic recombination).(Tonegawa, 1983) *IGKJ3* sits in an IGK cluster of genes located on the short arm of chromosome 2 whereby a 1800 kb segment encodes 76 IGK genes, 34 functional, 32 pseudogenes, 7 open reading frames and 3 unknown.(Barbie and Lefranc, 1998) Activation may reflect the body's immune reaction to SGC or aberrant activation.

Chemokine (C-X-C motif) ligand 9 (*CXCL9*) – chromosome 4q21.1

Chemokines are small signalling proteins (cytokines) released from cells and whose name originates from their ability to signal leucocyte traffic by inducing chemotaxis. In cancer they have competing roles and can have an anti-tumour effect by attracting immune cells or assist in tumour spread, invasion and proliferation.(Gorbachev and Fairchild, 2014) CXCL9 tends to act as an inflammatory chemoattractant rather than a homeostatic chemokine. It acts through CXCR3, a G Protein-coupled receptor (GPR); however, the

resultant effect is dependent on the receptor isoform. One effect is to activate the phosphatidylinositol 3-kinase (PI3K) and Protein Kinase B (Akt) pathway. Within tumours and its milieu, CXCL9 may have a dual role. The local tumour microenvironment harbours activated macrophages that release chemokines, especially CXCR3 receptor ligands, namely CXCL9/MIG and CXCL10/IP-10 that in turn attract CXCR3 expressing T-lymphocytes and natural killer cells. Nevertheless, tumour cells can also release chemokines and melanoma endothelial cells have been shown to secrete CXCL9, which subsequently assists in tumour migration across an endothelial monolayer.(Amatschek et al., 2011) CXCL10 was also overexpressed by log2FC 2.5 ($P<0.001$) compared to nodular log2FC 0.4, $P=0.397$.

Immunoglobulin kappa variable 1-5 (*IGKV1-5)* – chromosome 2p11.2

As mentioned, the immunoglobulin undergoes somatic recombination (V(D)J recombination) involving the variable subgene. Overexpression of IGKV1-5 is seen in chronic lymphocytic leukaemia and demonstrates a bias towards the use of this subgene rather than random somatic recombination. (Stamatopoulos et al., 2005). Having light chain bias (*IGKV1-5* and *IGKJ3)* prompts speculation that an over expressing functional kappa chain is important in SGC.

Chemokine (C-X-C Motif) Receptor 4 (*CXCR4*) – chromosome 2q22.1

In contrast to most chemokine receptors, CXCR4 is relatively monogamous to the chemokine CXCL12 (aka stromal derived factor 1, SDF-1). CXCL12 promotes blood vessel formation in the embryo and recruits endothelial progenitor cells in the adult. (Eman et al., 2014, Cavallero et al., 2015) In addition, CXCL12 is induced by proinflammatory stimuli. CXCR4 is overexpressed in many cancers, confers a poorer prognosis and linked to invasion and metastasis. (Ohtani et al., 2009, Chen and Zhong, 2015, Rave-Frank et al., 2015, Lombardi et al., 2013) Contradictory to this is a study in breast cancer demonstrating that CXCR4 correlates with prolonged disease free, metastasis free and overall survival, however, the same study states that higher levels also correlate with oestrogen receptor negative breast cancer which in its self-confers a poorer prognosis.(Dalm et al., 2015)

CXCR4 also binds ubiquitin, the small regulatory protein, and mediates an anti-inflammatory action. (Majetschak, 2011) Interestingly, ubiquitin specific peptidase 1

(*USP-1*) is significantly (P=0.002) overexpressed (expression 5.0) by log 1.2 in the pagetoid subtype only. Ubiquination is required for DNA repair and USP-1 removes ubiquitin (deubiquinates) thereby acting as a modulator of DNA repair. Furthermore, USP-1 deubiquitinates Proliferating Cell Nuclear Antigen (*PCNA*), Fanconi Anaemia Complementation Group D2 (*FANCD2*) and Group I (*FANCI*); all are important components of the Fanconi anaemia DNA repair pathway.

Another action of CXCR4 is to bind macrophage migration inhibitory factor (MIF). MIF suppresses the anti-inflammatory effects of glucocorticoids. In colorectal cancer, MIF production correlates with the metastatic potential and poorer prognosis. (He et al., 2009)

DBF4 Zinc Finger Pseudogene 1 (*DBF4P1*) –chromosome 10q23.1

Long non-coding (lnc) RNA have transcripts longer than 200 nucleotides and part of this family is *DBF4P1* at 1892 nucleotides long. Lnc RNA lack conservation across species in contrast to small RNA such as microRNA or snoRNA, but this assertion occurs when applying conservation rules used for protein coding sequencing (such as the retention of an open reading frame); different rules may be required for lncRNAs. (Mercer et al., 2009) Their functionality is still disputed; however, they have been shown to play a critical role in certain cancers. LncRNA Urothelial cancer associated-1 (*UCA1*), promotes cell cycle progression and proliferation in bladder cancer. (Xue et al., 2015) Higher expression of UCA1 has been shown in non-small cell lung cancer and related to histological grade and lymph node metastasis. (Wang et al., 2015b) Another example of functionality: the lncRNA HOXA cluster antisense RNA2 (*HOXA-AS2*), correlates with gastric cancer severity and potentially induces epigenetically silencing of P21 (a mediator of cell cycle G1 phase arrest) by binding to EZH2. (Xie et al., 2015, Huang and Yu, 2015) *DBF4P1* has also been linked to the mitotic cell cycle super pathway. (Institute, 2015)

(*RP11-44K6.2*) – chromosome 8

This is a short (102 bp) non-coding RNA that contains two exons, but termed a sense intronic transcript. It may be involved in methylation and histone modification.(Bhartiya, 2015)

Histone Gene Cluster 1, H2B Histone family, Member M (*HIST1H2BM*) – chromosome 6p22.1

See section 5.4.2. A similar gene was discovered using RNAseq on separate SGC samples, hence the detection of HIST1H2BM highlights that these histones may play a significant role in SGC, especially in light of the chromosomal instability depicted in Figure 5.10. Unopposed mitosis is the hallmark of cancer and HIST1H2BM has been validated as a breast cancer associated protein in vitro. (Choong et al., 2010)

Cyclin-Dependent Kinases Regulatory Subunit 2 *(CKS2*) – chromosome 9q22.2

CKS2 is involved in several stages of the cell cycle. It has been shown to downregulate TP53 and upregulate cell cycle regulators cyclin A, B1 and CDK.(Kang et al., 2009, Shen et al., 2013a) The latter two, cyclin B1-CDK protein kinase, are known as mitosis promoting factor and is an essential component to allow the cell to progress through the G_2 phase of the cell cycle.(Vassilev et al., 2006) *CSK2* has been shown to be linked with Y-box binding protein 1 (*YB-1*) whereby the silencing of *YB-1* downregulated *CKS2*.(Yu et al., 2010) Interestingly, *YB-1* was over expressed (9.1) by log 1.3 (P=0.008) in pagetoid only and may be due to this direct relationship with *CSK2*. Increased *CSK2* expression has been seen in hepatocellular, oesophageal, gastric, colorectal and prostate cancer.(Shen et al., 2010a, Wang et al., 2013b, Tanaka et al., 2011, Notterman et al., 2001, Lan et al., 2008) Furthermore, higher expression correlated with increased tumour size, poorer histological grade, worse TNM status and reduced survival.(Shen et al., 2010a, Tanaka et al., 2011) Hsa-miR-26a directly alters the expression of CSK2.(Lv et al., 2013)

Nucleolar and Spindle associated protein 1 (*NUSAP1*) – chr15q15.1

Microtubules are long, hollow structures making up the cytoskeleton of the cytoplasm and also found in the nucleus where they are involved in the mitotic spindle that separates chromosomes during mitosis and the meiotic spindle during gamete production. Microtubule-associated proteins (MAP), such as NUSAP1, regulate the movement of microtubules and can have both a stabilising or destabilising effect. *NUSAP1* is selectively expressed in proliferating cells at the transition of G2 to mitosis, localises to the central spindle microtubules, plus, in both excessive and depleted states, leads to defective chromosomal segregation.(Raemaekers et al., 2003) NUSAP1 is degraded by anaphase-promoting complex/cyclosome (APC/C)-Cdh1 E3 ubiquitin ligase.(Li et al., 2007) Positive regulators of *NUSAP1* transcription include v-myc avian

myelocytomatosis viral oncogene homolog (MYC), lin-9 DREAM MuvB core complex component (LIN9) and nuclear transcription factor Y, alpha (NF-YA).(Fujiwara et al., 2006, Reichert et al., 2010, Hussain et al., 2009) MYC was significantly (P<0.001), expressed (7.3) by log2FC 1.5 in the pagetoid subtype only. In addition, it has shown to promote expression by the loss of RB1 via the RB1/E2F transcription factor 1 (E2F1) axis. (Gulzar et al., 2013, Gordon et al., 2015) Moreover, WES SGC data highlights the loss of RB1 function. As mentioned, *RB1* is mutated in SGC and so augment its expression. Moreover, E2F1 is upregulated in both nodular (P=0.005, expression 4.9) and pagetoid (P=0.007, expression 4.9) by log2FC 1.3. A negative transcription regulator is forkhead box O1 (FOX01), a known tumour suppressor gene that is significantly (P=0.003) reduced in nodular by log2FC -1.1, but insignificantly (P=0.169) reduced (log2FC -0.4) in pagetoid.(Takano et al., 2007) CDK1 inhibits *NUSAP1* interaction with microtubules by phosphorylation and was found to be significantly (P=0.005) overexpressed (3.7) by log2FC 1.2 in pagetoid only. NUSAP1 has a role in prostate cancer, breast cancer and in aggressive subtypes of melanoma. (Gordon et al., 2015, Lauss et al., 2008, Ryu et al., 2007)

Chromosome 2 Open Reading Frame 47 (*C2orf47*) – chromosome 2q33.1

Open reading frames are potentially translatable DNA sequence that has a both a start and stop codon with a string of in-frame sense codons between them.(Andrews and Rothnagel, 2014) If they are less than 250 codons long they are deemed short ORF (sORF) and C2orf47 is a 291 amino acids long mitochondrial protein.(Hayden and Bosco, 2008) The function of C2orf47 is predicted to involve in RNA/nucleotide binding. (Delgado et al., 2014a) Hsa-miR-16 has been shown to reduce the expression of *C2orf47* and the miRNA study below showed this miRNA is significantly (P=0.03) over expressed (expression 9.1) by log2FC 4.8 in both the nodular and pagetoid subtype. Furthermore, C2orf47 has been implicated in cancer (gaining the term oncoORF) and is proportionally expressed in childhood acute lymphocytic leukaemia. (Yang et al., 2009) OncoORFs have also been associated with prostate, hepatic and brain cancer.(Delgado et al., 2014b)

Cyclin-Dependent Kinase Inhibitor 3 (*CDKN3*) - chromosome 14q22

CDKN3 has to dual role as a protein phosphatase inhibitor in the control of the cell cycle. Firstly, it combines with and dephosphorylates Cyclin-Dependent Kinase 2 (CDK2), hence prevents its activation, which in turn reduces the aforementioned RB1/E2F axis

repression of transcription of genes needed for the S phase of the cell cycle.(Yeh et al., 2003, Niculescu et al., 2004) Secondly, it promotes the progression of the cell cycle by desensitising p21 action when combined with TP53 and E3 ubiquitin-protein ligase Mdm2 (MDM2) which in turn reduces TP53 gene target production.(Demetrick et al., 1995) It acts as a survival biomarker and inactivation decreases cell proliferation in cervical and ovarian cancer.(Barron et al., 2015, Li et al., 2014b, Zhang et al., 2015b) It is also overexpressed in lung adenocarcinoma, breast and colorectal carcinoma. (Yang and Sun, 2015, Taylor et al., 2010, Xing et al., 2012, MacDermed et al., 2010)

Structural Maintenance of chromosomes Flexible Hinge Domain Containing 1 (*SMCHD1*) – Chromosome 18p11.32

Structural maintenance of chromosome (SMC) family of proteins have a dual role in the mitotic segregation of chromosomes, namely chromosome condensation and sister chromatid cohesion.(Ball Jr and Yokomori, 2001) In addition, they play a role in DNA recombination and repair.(Hopfner et al., 2000) SMCDH1 is recruited to sites of DNA damage (along with Ku80 and RAD51) and its deficiency reduces cell survival in response to DNA damage.(Coker and Brockdorff, 2014) *SMCHD1* also regulates X chromosome lyonisation (through maintenance of hypermethylation of CpG islands), carries out telomere silencing, and regulates a subset of monoallelic expressing gene clusters found on autosomes.(Mould et al., 2013, Grolimund et al., 2013) Mutations in *SMCHD1* can lead to facioscapulohumeral muscular dystrophy type 2 (FSHD2) and in the physiological state, is implicated in silencing of Double Homeobox 4 (*DUX4*), whose encoded protein is overexpressed in FSHD2.(Daxinger et al., 2015) It has been suggested that it is a tumour suppressor gene in some leukaemia and lymphomas where it is often down regulated.(Leong et al., 2013) In addition, it may have a role to play in the perturbations of the inactive X chromosome seen in cancer.(Chaligne and Heard, 2014)

Lacritin (LACRT) gene – chromosome 12q13.2

The lacrimal gland, accessory lacrimal glands of Wolfring that are found throughout the conjunctiva, and meibomian glands express high levels of lacritin that is in the secretion of tears.(Ubels et al., 2012, Liu et al., 2011b) LACRT was significantly (P=0.009) downregulated by log2FC -4.0 in the pagetoid subtype and less so in nodular (P=0.01, log2FC -3.1).

5.12.3 Shared nodular and pagetoid mRNA

Thymidylate Synthetase *(TYMS*) – chromosome 18p11.32

Deoxyuridine 5'-monophosphate (deoxyuracilmonophosphate or dUMP) is converted to deoxythymidine monophosphate (dTMP) using thymidylate synthase which uses a methyl group from N^5,N^{10}-methylene tetrahydrofolate (THF). This is an essential step for DNA replication and repair. Chemotherapy agent 5-fluorouracil inhibits thymidylate synthase and gets misincorporated into nucleic acid (both DNA and RNA) causing failed replication.(Longley et al., 2003) TYMS overexpression is seen in aggressive forms of prostate cancer, lung, breast, gastric, colorectal and renal cell carcinoma.(Burdelski et al., 2015, Kotoula et al., 2012, Pestalozzi et al., 1997, Formentini et al., 2004, Popat et al., 2004) The exact role is not clear, but it has been shown to promote malignant phenotype in excess and lead to drug resistance. (Rahman et al., 2004, Ahn et al., 2015) Moreover, the may be a mechanistic link between TYMS overexpression and chromosomal abnormalities in prostate cancer and it is possible there is a link too with SGC which is chromosomally unstable.(Burdelski et al., 2015)

Histone Cluster 1, H2bd (*HIST1H2BD*) – chromosome 6p22.2

Histone Cluster 1, H2ae (*HIST1H2AE*) – chromosome 6p 22.2

Histone Cluster 1, H2bm (*HIST1H2BM*) – chromosome 6p22.1

All three are up regulated significantly in both subtypes reinforces the possibility they are important driver gene for both and may represent an early 'trunk' derangement in SGC.

Small Nucleolar RNA Host Gene 1 (*SNHG1*) - chromosome11q12.3

Activated TP53 can act to repress or activate target genes. More recent evidence has shown that it can regulate miRNA, lncRNA and snoRNA including the repression of SNHG1.(Yu et al., 2015a) The exact role of this gene is not clear, however, snoRNA can be processed into smaller miRNA sometimes termed sno-miRNA or sno-derived small RNAs (sdRNAs).(Taft et al., 2009) SNHG1 can be processed into SNORD25, SNORD28 and sno-miR-28 and is negatively regulated by p53. (Yu et al., 2015a) However, SNHG1 is induced in irradiated lymphoblastoid cell lines which were TP53 positive and negative, suggesting an alternate regulator after a genotoxic insult.(Chaudhry, 2013) Its role as an oncogene is suggested by its interaction with TP53 along with the overexpression expression of SNHG1 is seen in breast cancer, gastric cancer and lung cancer. (Cao et al.,

2013, Yu et al., 2015a, You et al., 2014) Loss of TP53, as suggested by the WES data may allow for uncontrolled SNHG1 expression in SGC.

Other important shared genes

Notch 2 N-Terminal Like (*NOTCH2NL*)- chromosome 1q21.2

This was the highest expressed (over 9) gene in both subtypes although the log2FC scores were only slightly more than control at log2FC 1.5 (P=0.001), log2FC 1.2 (p=0.007) for nodular and pagetoid respectively. The log2FC score was not particularly high and so it did not enter the top differential expressed shared genes, but its role may still be important. Notch signalling is important for physically adjacent cell communication and is involved in a raft of processes including cell fate decisions. These decisions include stem cell/progenitor cell pool and differentiating cell lineages. Notch interaction with other pathways has been shown to be important in cancer transformation; however, there are some instances where it can be adverse to tumour progression.(Grishina, 2015, Miele et al., 2006, van Es and Clevers, 2005) There are four Notch receptors and Notch2 has shown to confer aggressive behaviour in hepatocellular carcinoma cells.(Hayashi et al., 2015) Mutations have been seen in *NOTCH2NL* mutations are seen in anaplastic astrocytoma (WHO grade III), a more aggressive form of brain tumour. (Killela et al., 2014)

Frizzled Class Receptor 6 (*FZD6*) – chromosome 8q22.3

Frizzled receptors are part of the Wnt signalling pathway. *FZD6* was overexpressed by log2FC 2.1 (expression 6.4, P<0.001), log2FC 1.9 (expression 6.2, P<0.001) in nodular and pagetoid respectively. Wnt4 binds to FZD6 to act as a negative regulator of canonical beta-catenin Wnt signalling via a non-canonical route: transforming growth factor-beta-activated kinase (TAK1)-NEMO-like-kinase (NLK) pathway.(Golan et al., 2004, Lyons et al., 2004) The receptor also binds secreted-frizzled related protein 1 (SFRP1), a known tumour suppressor that acts as a negative feedback response to beta-cantenin/TCF activity.(Bafico et al., 1999) Furthermore, Hedgehog-Gli signalling induces SFRP1 to keep differentiating epithelial cells away from the effects of canonical Wnt signalling.(Katoh and Katoh, 2006) Despite the suppressive actions of the receptor, its overexpression is seen in neuroblastoma and squamous cell carcinoma cancer, but not in benign skin hyperplasia.(Haider et al., 2006, Cantilena et al., 2011) FZD6 was associated with aggressive neuroblastoma cells with stem cell features (expressing Twist1 and

Notch1) and chemoresistance. (Cantilena et al., 2011) Hsa-miR-199 negatively regulates FZD6 and this miRNA is significantly (p=0.002) reduced by log2FC -2.3 in pagetoid only.

5.13 SGC microRNA subtype analysis

5.13.1 Nodular SGC microRNA

Hsa-miR-150 – chromosome19q13.33

Hsa-miR-150 plays a role in haematopoiesis and differentiation of C and T lymphocytes, the later in conjunction with hsa-mir-155.(Kozomara and Griffiths-Jones, 2014) Increased expression is correlated with a poorer survival in patients with prostate cancer.(Dezhong et al., 2015) It promotes growth and malignant behaviour in breast cancer by suppressing the pro-apoptotic receptor, Purinergic Receptor P2X, Ligand Gated Ion Channel, 7 (*P2X7*) expression.(Huang et al., 2013) In contrast, it has shown to suppress Mucin 4, Cell Surface Associated (*MUC4*) by binding to the 3' UTR with an associated decrease in HER2 expression that suppresses growth and malignant behaviour in pancreatic cancer cells, providing evidence for a role as a tumour suppressor miRNA.(Srivastava et al., 2011) Tumour suppressor action of hsa-miR-150 has also been noted in epithelial ovarian cancer and shown to reduce invasion and metastasis by suppressing the transcriptional repressor, Zinc Finger E-Box Binding Homeobox 1 (*ZEB1*). (Jin et al., 2014) Interestingly, *ZEB1* is significantly (P=0.004) downregulated by log2FC -1.1 (expression 4.6) compared to control and pagetoid, which correlates with hsa-miR-150 expression in these tissues suggesting that it is playing a tumour suppressor role in nodular SGC through *ZEB1* suppression. Further tumour suppressive roles have been noted in colorectal and hepatocellular cancer by targeting V-Myb Avian Myeloblastosis Viral Oncogene Homolog (*MYB*), Matrix Metallopeptidase 14 (Membrane-Inserted) (*MMP14*).(Feng et al., 2014, Liang et al., 2007) Moreover, MYB, MMP14, HER2, P2X7 and MUC4 are not differentially expressed in SGC. The tumour suppressive action may help nodular SGC to behave in a less aggressive fashion compared to its aggressive pagetoid counterpart.

Hsa-miR-142-3p – chromosome 17q22

This is a dominantly expressed miRNA in the adult haemopoetic system including megakaryocyte maturation and maintaining normal actin cytoskeletal dynamics. (Chapnik et al., 2014) It acts as a tumour suppressor and represses Transforming Growth Factor, Beta Receptor 1 (*TGFRB1)* in non-small cell lung cancer; Prominin 1 (*PROM1*), ATP-Binding Cassette, Sub-Family G (WHITE), Member 2 (Junior Blood Group) (*ABCG2*), and Leucine-Rich Repeat Containing G Protein-Coupled Receptor 5 (*LRG5)* in colon cancer cells, and Frizzled Class Receptor 7 (*FZD7*) in cervical cancer.(Lei et al., 2014, Shen et al., 2013b, Deng et al., 2015)

Hsa-miR-143 – chromosome 5q13.3

This miRNA is highly conserved and thought to be involved in cardiac morphogenesis. In cancer, it is suppressed in metastatic prostate cancer and when re-expressed, prevents cell proliferation.(Coarfa et al., 2015) It is also downregulated in gastric cancer and when over expressed, is potent at autophagy via GABA (A) Receptor-Associated Protein Like 1 (*GABARAPL1*).(Du et al., 2015) Moreover, hsa-miR-143 is inversely correlated to Inhibin, Beta A (INHBA) in oral squamous cell carcinoma, whose high levels of protein activin A are correlated with lymph node metastasis, tumour differentiation and poor survival.(Bufalino et al., 2015) Further tumour suppressor actions are seen in prostatic cancer where it targets hexokinse 2 and it is down regulated in breast cancer.(Zhou et al., 2015, Chang et al., 2015)

Hsa-miR-548g – chromosome 4q31.23

The family of miR-548 are found across all chromosomes, but poorly conserved evolutionary and primate specific.(Liang et al., 2012) The sequence of this miRNA is found within a tumour suppressor gene, Fragile Histidine Triad *(FHIT)* that may suggest that it has a related role to play.(Hu et al., 2014) Downregulation of miR-548-3p is seen in breast cancer cells, and when it was over expressed, a decrease in proliferation and concomitant rise in apoptosis noted. (Shi et al., 2015b) Loss of this potential TSG may confer an advantage to SGC.

Hsa-miR-603 – chromosome 10p12.2

Over expression is seen in glioma tissue and it directly targets WNT Inhibitory Factor 1 (*WIF1*) and Catenin, Beta Interacting Protein 1 (*CTNNBIP1*); the former activates the

Wnt/β-catenin pathway.(Guo et al., 2015) Excess is also seen in pancreatic cancer.(Yu et al., 2012) It is found to be downregulated in laryngeal squamous cell carcinoma where it was shown to inhibit *E2F1*, a controller of tumour suppressor proteins.(Ayaz et al., 2013) It targets *CCND1* and *CCND2* expression by acting as a tumour suppressor gene in thyroid carcinoma.(Mussnich et al., 2013) Loss of this potential TSG may confer an advantage in nodular SGC.

5.13.2 Pagetoid SGC microRNA

Hsa-miR-1308 - chromosome

Although, this apparent miRNA has been found to be raised in multiple myeloma, it's actually a 5' cleaved fragment of tRNA.(Jones et al., 2012)

Hsa-miR-205 – chromosome 1 q32.2

Increased levels of this miR-205 prevents TGFB1 induced epithelial to mesenchymal transition (EMT) in vitro and specifically inhibits the cadherin-1 (CDH1) transcriptional repressors *ZEB1* and *ZEB2*.(Gregory et al., 2008) Its action on ZEB1 also radiosensitises cells by preventing DNA repair and is down regulated in radioresistant breast cancer cells.(Feng et al., 2014) SGC does seem to be radiosensitive with good survival rates at 5 years in a small, 13 patient case series and radiotherapy remains an option in those who are too unwell or refuse surgical intervention.(Hata et al., 2012) Downregulation seems be an important step for metastatic tumours, however, raised levels of hsa-miR-205 promoted vascular endothelial growth factor (VEGF) invasion of ovarian cancer cells by targeting Ezrin (*EZR*) and lamina A/C (*LMNA*).(Li et al., 2015a) Higher levels are associated with adverse clinical outcome in bladder cancer through its action on ZEB1 (by maintaining an epithelial phenotype) and ultimately being controlled by tumour protein 63.(Tran et al., 2013) Although this goes against the logic that EMT is required for metastasis, EMT plasticity is likely to be required for cells in distant sites to proliferate in their new locations (i.e. undergo MET). ZEB1 is down regulated in nodular -1.1 log2FC (P=0.004, expression 3.5) but there is non-significant change in the pagetoid at -0.3 log2FC (p=0.387, expression 4.3). Upregulation of hsa-miR-205 occurs in NSCLC and directly represses PTEN and PHLPP2 expression, which in turn activates the AKT/FOXO3a and AKT/mTOR signalling pathways respectively.(Cai et al., 2013) An

interaction between EZH2, MALAT1 and hsa-mir-205 to control b-catenin through EZH2 occurs in renal cell carcinoma.(Hirata et al., 2015) EZH2 is over expressed in both subtypes, 1.4 log2FC (P=0.01, expression 5.4) nodular and pagetoid 1.6 log2FC (P=0.004, expression 5.7).

Hsa-miR-200a – chromosome 1 p36.33

The miR-200 family are involved in EMT as mentioned above. In addition, hsa-miR-200a has a physiological role in dendritic cell maturation and activation in unison with growth hormone via the Kelch-Like ECH-Associated Protein 1/ Nuclear factor erythroid 2-related factor 2 (Keap1/Nrf2) pathway, a critical pathway that protects cells from oxidative and electrophilic stress.(Liu et al., 2015d) Activation of this pathway through hsa-miR-200a to protect cancer cells from oxidative stress also occurs in oesophageal squamous cell carcinoma.(Liu et al., 2015c) It works with hsa-miR-141 to target mitogen-activated protein kinase 14 (MAPK14) to enhance the oxidative stress tumour growth response in ovarian cancer.(Mateescu et al., 2011, Zhu and Gao, 2014) Interestingly, hsa-miR-141 is significantly (p=0.009) upregulated by log2FC 2.6 (expression 7.1) compared to control and nodular subtype. In breast cancer, hsa-miR-200a has been shown to suppress Gap junction alpha-1 protein (GJA1) that aids in metastatic spread. Other tumour suppressor actions include inhibiting tumour growth in neuroblastoma cells by targeting Amyloid Beta (A4) Precursor Protein (APP).(Gao et al., 2014)

Hsa-miR-1260 – chromosome 14q24.3

Hsa-miR-1260 probably has a role in the PI3K/Akt pathway in hepatocellular cancer cells by targeting G1/S-specific cyclin-D1 (CCND1).(Yan et al., 2013) Overexpression is seen in cutaneous malignant melanoma, oestrogen receptor positive breast cancer and downregulation is seen in gastric cancer.(Sand et al., 2013, Park et al., 2014, Ma et al., 2013) The latter may have a reciprocal relationship with hsa-miR-29 in gastric cancer and this is significantly (P=0.0002) down-regulated (Log2FC -1.6, expression 9.8) in the pagetoid subtype, thus supporting a reciprocal relationship.

Hsa-miR-199a-3p+miR-199b-3p – chromosome 19p13.2 + 9q34.11 respectively

This miRNA acts as a tumour suppressor gene in several cancers including ovarian carcinoma, osteosarcoma, colorectal cancer and thyroid cancer .(Kozomara and Griffiths-Jones, 2014, Tian et al., 2014, Han et al., 2014, Minna et al., 2014) Furthermore, it is

upregulated at a post-transcriptional level by TP53.(Wang et al., 2012a) Hsa-miR-199a+3p has been shown to target CD44 glycoprotein (an oncoprotein) that aids in osteosarcoma cell adhesion and migration, with increased CD44 expression contributing to the aggressive nature of the tumour.(Gao et al., 2015) This inverse association has also been noted in hepatocellular carcinoma.(Henry et al., 2010) CD44 was found to be significantly (P=0.005, expression 7.8) overexpressed (log2FC 1.2) in pagetoid subtype only, which would support the reciprocal association with hsa-miR-199a+3p. It has also been shown to target AXL Receptor Tyrosine Kinase (*AXL*) which in turn regulates invasion and metastasis through the PI3K/AKT pathway.(Tian et al., 2014)

5.13.3 Comparison of nodular and pagetoid SGC microRNA

Hsa-miR-16 – chromosome 13q14.3

In over half of chronic lymphocytic leukaemia cases there is a loss of chromosome 13q where this miRNA resides. Subsequent to this study, hsa-miR-16, along with an associated cluster of miRNA, are found to target cyclin D1, MCL1, and Bcl-2.(Pekarsky and Croce, 2015) The latter being over expressed in almost all CLL cases and is a direct target of hsa-miR-16 and -15. Bcl-2 is non-significantly (P=0.7) downregulated (log2FC -0.2, expression 4.0) in the nodular subtype, but still overexpressed (log2FC 1.5, expression 5.6, P=0.009) in pagetoid. Hsa-miR-16 has anti-angiogenic properties by targeting VEGF and fibroblastic Fibroblast Growth Factor Receptor 1 (FGFR1). (Caporali and Emanueli, 2011)

Hsa- miR-9 – chromosome 1 q23.1

Physiologically, this regulates neural differentiation in the central nervous system by targets FoxG1, Hes1 or Tlx and involved in apoptosis.(Coolen et al., 2013, Yuva-Aydemir et al., 2011) In breast cancer, hsa-miR-9 excess leads to increased cell motility and invasiveness by targeting CDH1; its expression is activated by MYC and MYCN.(Ma et al., 2010)

.

Hsa- miR-34a – chromosome 1p36.22

This miRNA forms part of the TP53 suppressor network by modulating TP53 targets and forms a positive feedback loop via sirtuin 1 (SIRT1).(Lai et al., 2012) It has many suppressive effects in several different cancers including E2F Transcription Factor 3

(E2F3) in neuroblastoma plus colorectal cancer, MET Proto-Oncogene, Receptor Tyrosine Kinase (MET) in hepatocellular carcinoma, CD44 in prostate cancer, AXL in ovarian cancer and G1/S-specific cyclin-E1 (CCNE) in lung cancer.(Cho, 2007, Tazawa et al., 2007, Li et al., 2009, Li and Lam, 2015, Li et al., 2015c, Han et al., 2015) Nevertheless, in the presence of over expressed MYC, hsa-miR-34a behaves differently and as mentioned earlier, MYC is over expressed in the pagetoid subtype. Specifically, in myc-driven tumours, hsa-miR-34a improves cell survival based on its ability to reduce TP53 levels, in a myc mediated and dependent fashion, hence is helping the cancer cells to survive.(Sotillo et al., 2011)

Hsa-miR-126-chromosome 9q34.3

Hsa-miR-126 is expressed in epithelial cells, particular vascular endothelial cells where it plays a role in pro-angiogenesis by inhibiting Sprouty-Related, EVH1 Domain Containing 1 (SPRED1).(Wang et al., 2008) It acts as a TSG by targeting V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog (CRK) in gastric cancer, ADAM9 in breast cancer, PAK4 in ovarian cancer, RhoA/ROCK signalling in colorectal cancer and SLC7A5 in small cell lung cancer.(Feng et al., 2010, Wang et al., 2015a, Luo et al., 2015, Li et al., 2013) Loss of this TSG confers a benefit for SGC. In pancreatic carcinoma, loss of hsa-miR-126 is crucial for progression via ADAM9. Down regulation in oral squamous cell carcinoma induces angiogenesis and lymphangiogenesis, conferring a poorer prognosis.(Sasahira et al., 2012)

Hsa- miR-619 – chromosome 12q24.11

Little data has been published on this miRNA. It has been associated with coronary heart disease.(Hou et al., 2014) Predicted gene targets include MAPK9.

Hsa-miR-21 – chromosome 17q23

Oncomir 21 was present in both nodular (Log2FC 2.6; exp 9.7; $P<0.01$) and pagetoid (Log2FC 2.3; exp 9.7; $P<0.02$). It modulates PTEN and BCL2.

CONCLUSION

5.14 Summary of results

Overall mutational burden of SGC was low and more in keeping with a visceral tumour that only contains a few hundred variants. This was unexpected; as an eyelid lesion it was thought that it might also be a UV induced tumour like BCC. Furthermore, it was thought that its close proximity to the skin with penetration through the skin on occasion along with spread along the visible bulbar conjunctiva would allow for the collection of UV type mutations. The later may explain some of the more aggressive SGC samples accumulating more UV type mutations. There was a large proportion of G>T (C>A) changes which may reflect a tobacco type or a defective DNA mismatch repair signature.

Analysis of specific mutations identified five shared mutations between the two algorithms including two, well known tumour suppressor genes TP53 and RB1. Intogen analysis picked the dynein family that was ubiquitously affected in all samples. DNAH14 has recently been describes as a driver mutation in ovarian cancer (Braun, 2013). They are thought to play a role in chromatid segregation and may be involved in the chromosomal instability.

WES revealed four pathways significantly mutated in all five samples, namely the cell cycle, pancreatic cancer, Wnt signalling pathway and dopaminergic synapse. Pancreatic cancer is one of the most aggressive cancers with a poor 5-year survival rate and behaves in a similar fashion. Locally invasive behaviour may be explained by mutations found within the ubiquitin mediated proteolysis pathway, which is also differentially expressed.

Chromosomal instability was highlighted with areas of amplification at chromosome 8 where the MYC gene resides and loss of the short arm of chromosome 17 where TP53 sits. SGC therefore, comprises genomic instability rather than mutational instability. Further pointers come from the RNA analysis with both modalities underlining the over-expression of SYCP2, part of a protein complex required for the joining of homologous chromosomes during planned DSB.

RNAseq Pathway showed upregulation of mismatch repair, nucleotide excision repair, and homologous recombination supporting the chromosomal level of change. SGC often

spreads in a pagetoid fashion with skip lesion and invades with ease. Downregulation of cell adhesion, extracellular matrix organisation and collagen formation would aid with this locally invasive behaviour. Correlation of WES and RNAseq data demonstrated upregulated 'cell cycle', 'ubiquitin mediated proteolysis' and 'Wnt signalling' pathways. These pathways together will aid tumour growth and assist in extracellular matrix breakdown for local spread.

Subtype analysis of pagetoid and nodular SGC revealed the histone gene cluster family as important to both, again supporting chromosomal instability. Further chromosome genes included NUSAP1 and SMCHD1 in the pagetoid type along with important chemokines that may assist is pervasive spread.

Oncomir hsa-miR-21 was found in both subtypes and targets tumour suppressor genes PTEN and PDCD4. Loss of hsa-miR-199a occurs in the pagetoid subtype and this was confirmed using RT-qPCR. This allows uncontrolled oncoprotein CD44 activity that in turn promotes invasion plus metastasis. Furthermore, it targets AXL, which promotes invasion and metastasis, thus confirming that hsa-miR-199a is a tumour suppressor. Another pagetoid specific miRNA was hsa-miR-205a whose over-expression is thought to promote local invasion, and although an inhibitor of EMT, may push distant metastasis into MET, thus allowing them to colonise their new location. Furthermore, it controls b-catenin (Wnt signalling is upregulated) through an interaction with EZH2 and MALAT1. Nodular specific hsa-miR-150 can suppress pro-apoptotic factors, but has also been shown to behave in a tumour suppressive manner and is thought to act via ZEB1 suppression, which is downregulated in nodular SGC only.

Protein expression of HIST1H2BD showed increased staining in both subtypes supporting aforementioned data on chromosomal instability. Surprisingly, protein analysis also demonstrated the expression of the canonical Hh pathway, namely PTCH1, SMO, Gli1 and Gli2 in SGC. Moreover, the expression patterns are beyond normal reactivation levels, suggesting aberrant signalling. Hedgehog pathway activation occurs in periocular SGC, in both the tumour and surrounding stromal tissue at a higher level than nodBCC, a known Hh driven tumour. The timing for any treatment modification of the Hh pathway would be dependent on its specific role in the pathogenesis of SGC, and further studies are required to determine this.

5.15 Potential treatment targets

Activating tumour suppressor genes (TSG)

Blocking an overexpressed kinase or an oncogene has been the mainstay of pharmacological treatment in cancer. Reactivating a tumour suppressor gene is an onerous task, but restoration of wildtype TP53 has improved the radio or chemosensitivity of a tumour. The concept of priming the body's defences to a kill the tumour may be more beneficial route in the future rather than a pan-toxic drug that targets dividing cells. Both TP53 and RB1 are mutated in SGC and represent an attractive target. How to restore or produce wildtype TSGs requires a different approach and ultimately needs to be personalised to the tumour in question. Activation of TSG regulators is one route: MDM2 is a negative regulator of TP53 and inhibiting their interaction results in an increase in TP53. Nutlin was the first MDM2 molecule and has undergone clinical trials.(Khoo et al., 2014) Second generation have already been developed that are less likely to succumb to tumour resistance and are undergoing clinical trials. (Morris and Chan, 2015) Presumably, some wildtype TP53 is required.

Matrix Metalloproteinase 9 (MMP9) inhibitors

Extracellular matrix can be modified by metalloproteinases, which also modify inflammation and growth. MMP9 is especially associated with pathological inflammation in vascular disease, inflammatory bowel disease and cancer.(Rodriguez et al., 2010) REPIN1 was picked up as a driver mutation using two algorithms and regulates MMP9 production. Non-selective MMP inhibitors have been used in humans with mixed results and significant side effects. Being selective for MMP9 has been a focus since the initial results of the non-selective inhibitors and two have been developed, the antibodies REGA-3G12 and GS-5745, the latter being in phase 1 clinical trials for solid tumours (ClinicalTrials.gov identifiers NCT01831427, NCT02077465, and NCT01803282).(Marshall et al., 2015)

Histone deacetylase 1 (HDAC1) inhibitor

See section 4.7

Pre-MRNA Processing Factor 31 (*PRPF31*) - non-viral gene therapy

Small molecules or nanoparticles have been developed as a delivery system as a non-viral gene therapy to restore PRPF31 in autosomal dominant RP.(Pensado et al., 2016) Injection into the subretinal space in Prpf31A216P/+ mice restored visual acuity and highlights that restoration of this gene is possible. However, the expression of PRPF31 was neither up nor down, so it is not clear if it is a loss or gain of function and restoration would only be beneficial if it was a TSG.

S100 Calcium Binding Protein P (*S100P*)

S100P has been identified in a vast number of cancers and normally overexpressed. S100P interacts with RAGE stimulate cell growth and this interaction is blocked using Cromolyn and a 5-methyl analogue which is more potent.(Arumugam et al., 2013) Antibodies to S100P have shown promise in combination with chemotherapy agents in pancreatic cancer.(Dakhel et al., 2014) A novel method of reducing S100p is by incorporating an antisense mRNA within a retrovirus to target cancer cells.(Prica et al., 2016)

Histone Cluster 1, H4h (*HIST1H4H*) – DNA methylating agent

Many of the histone family of proteins are altered in SGC. Temozolomide is an alkylating agent used in the most aggressive forms of brain tumour and glioblastoma multiforme. It has also noted that it can methylate histone proteins and this is thought to be part of its anti-cancer action.(Wang et al., 2016a)

Epithelial Cell Adhesion Molecule (EPCAM) inhibitor

Epithelial cancers often express the epithelial cell adhesion molecule and represent an attractive target. It is not only an adhesion molecule, but also involved in signalling, migration, proliferation and differentiation.(Ohashi et al., 2016) Overexpression correlates with aggressive behaviour and thought to assist in invasiveness plus render the tumour chemo-resistant. Catumaxomab is an antibody initially given via the intraperitoneal route for malignant ascites was noted to benefit tumours outside of the treatment zone suggesting an overall anti-cancer benefit. A phase 1 trial of repeated intravenous (rather than intraperitoneal) administration of 7µg in solid epithelial tumours (mainly colorectal cancer) demonstrated with chills, pyrexia and hepatotoxicity.(Mau-Sorensen et al., 2015) Although not part of the phase 1 study, none of the 16 patients

demonstrated an effective reduction in tumour burden. Rather than a treatment modality, it could be used as a monitoring of tumour burden or recurrence as newer technologies have been developed to detect very small amounts of EPCAM positive tumour cells within the blood.(Andree et al., 2016)

Thymidylate Synthetase (TYMS)

TYMS was one of four shared, overexpressed genes in both nodular and pagetoid SGC. It is an enzyme involved in DNA synthesis and has been a target for chemotherapy agent 5-fluorouracil for some time. Pemetrexed is a newer anti-folate, which is active in a variety of solid tumours.(Adjei, 2001) Variants in TYMS have been associated with chemoresistance to 5FU, especially when in excess. Incidentally, platinum based agents seem to be more effective despite not traditionally known as an anti-folate agent.(Kotoula et al., 2012)

Pathway inhibition

Wnt signalling inhibition

Both WES and RNAseq data support the overexpression of the Wnt signalling pathway in SGC. Both subtypes were shown to overexpress FZD6. Moreover, loss of hsa-miR-199a, which targets FZD6 in pagetoid, compounds the situation. See section 3.8 for treatment possibilities.

Hedgehog pathway inhibition

Activation of the Hh pathway in SGC raises the possibility of medical treatment for these tumours, and potentially avoiding exenteration. Attenuation of the Hh pathway has been shown to be successful by using SMO inhibitors, with Vismodegib presently being the most widely used compound in clinical trials and used for locally advanced or metastatic BCC.(LoRusso et al., 2011, Demirci et al., 2015) Furthermore, neoadjuvant treatment with Vismodegib to reduce tumour size in BCC has been successful in making tumours more amenable to surgery.(Ally et al., 2014) Vismodegib has been trialled in a variety of cancers, including pancreatic, ovarian cancer, breast and prostate cancers and other SMO inhibitors are emerging such as Sonidegib and BMS-833923.(Queiroz et al., 2012) It is possible these agents could be efficacious in advanced or metastatic SGC and if such a neoadjuvant response could be replicated it may allow the preservation of sight by averting the need for a blinding exenteration.

CHAPTER SIX

6 GENERAL DISCUSSION AND CONCLUSION

6.1 Overview of study

Cancer is an increasing burden on humanity with NMSC being the most prevalent. BCC, the commonest NMSC, within the periocular region exhibits a high risk of recurrence, especially the morphoeic subtype. In contrast, SGC is a rare life-threatening cancer that has a predilection for the periocular region and often require mutilating surgery to save life.

BCCs within the H zone, including periocular, have a higher risk of recurrence, behave more aggressively and those around the eye area are a risk to sight.(Mosterd et al., 2008) Compounding this is a histological diagnosis of morphoeic BCC (mBCC), which is a high risk factor for recurrence and often requires Mohs micrographic surgery to ensure complete histological resection.(Malhotra et al., 2004) SGC has two broad macroscopic presentations, namely nodular or pagetoid. Misdiagnosis is common, often labelled as a benign chalazion in its nodular form or conjunctivitis in the conjunctival, pagetoid type. Thus, it is described as a masquerading lesion with delays to diagnosis often greater than a year. Pagetoid expansion of the skin or conjunctiva carries a higher risk of orbital exenteration.(Chao et al., 2001) Despite aggressive treatment there is a high recurrence rate and metastasis often occurs via the lymphatics to the cervical lymph nodes.(Buitrago and Joseph, 2008) Mortality rates were once very high at 29%, but there is a significant range depending on which treatment modality is employed and can be as good as 6% in a centre of excellence.(Doxanas and Green, 1984, Shields et al., 2004, Ni, 1982).

Whole exome sequencing (WES) of 10 periocular mBCC were compared to 10 nodBCC; 5 periocular and 5 below the head to allow for histological and site specific evaluation. BCCs in the periocular region were hypothesised to have a greater tumour burden and higher UV signature than BCC elsewhere (elseBCC) with mBCC being the most deranged. Overall burden was less in mBCC than nodBCC, mutational signature was not different and locality had no effect. The confounding factor is that normal skin cells

contain thousands of mutations including those in damaging genes such as *NOTCH1* and *TP53* in up to 21% and 5% of the time respectively. (Martincorena et al., 2015)

In contrast, WES of SGC revealed a visceral tumour like mutational burden suggesting that is not originating from any form of skin cell. Nevertheless, if mutational burden is not relevant to pathogenesis, then the specific types of mutations present within the tumour are what define the tumour's aggressiveness. Two tumour suppressor genes were found to be associated with SGC, along with REPIN1 that regulates stromal metalloproteinase MMP9 and underline its locally invasive behaviour. By identifying those driver genes specific to the tumour such as *HECTD4* and *FLNB* in mBCC, it may give the clinician the ability to predict outcome. For example, if there is a tumour that looks not too dysplastic on histology but harbours the mBCC mutations then perhaps these are the patients that need closer follow-up. The functional consequence of these genetic changes needs to be ascertained and the use of 3D BCC in vitro models or organotypics should allow this to be done (see future work, section 6.2.1). Adding these genes to a known list of cancer associated genes and testing them in patients could help to stratify risk and lead to a more personalised approach to cancer care (see future work, section 6.2.3). Moreover, the presence or absence of these genes may be able to help with the prognosis and determine who require regular follow up and who may be discharged (see section 6.2.3).

Perhaps a more sensible approach is to look at a family of genes rather than a specific gene or variant, for example, the dynein or histone family in SGC. It is possible that they act in a similar fashion so that any one of the genes will result in chromosomal instability, nevertheless modelling these changes would help to elucidate their true function. We know that genomic instability is not solely mutational and chromosomal instability may play a large role in cancer behaviour. Large number of genes can be lost or amplified in a single event and arguably more detrimental to the cell than a single base change. Karyotyping may have gone slightly out of fashion, but when NGS becomes better at looking at the chromosome, then a clearer picture on overall genomic instability will be discerned. Currently NGS is a weak at looking chromosomes especially with difficult to detect abnormalities such as inversions or large deletions. Karyotyping is still looked at in clinical practice in terms of prognosis such as the presence of trisomy 12 within the tumour being a poorer prognosis in retinoblastoma.(AbdelSalam et al., 2008)

Presence of the gene or mutation is only the first step and its subsequent expression (loss or gain of function), is important to determine. For example, loss of hsa-miR-199a expression may indicate pagetoid behaviour and poorer prognosis in SGC. Expression of a whole pathway rather than just one gene may be more relevant to cancer behaviour. Traditionally we look at the fold change of a gene such as 4.5x increase in hsa-miR-205-5p in pagetoid SGC. However, a small increase or decrease all 140 genes of a pathway is likely to have more profound effect than just a 4.5 fold change in one gene. Detecting a loss of immune pathways such has 'Natural killer cell mediated cytotoxicity' and 'Fc Epsilon RI signalling pathway' may indicate a poor prognosis in BCC.

6.2 Future work

6.2.1 Organotypics and testing therapeutics

An *in-vitro* cell model of BCC has been developed in our lab using a the Matrigel 3D organotypic system as a model for skin: The dermis is formed using Matrigel, human fibroblasts and collagen. Subsequently, human keratinocytes are seeding on top of the dermal model and cultured using an air-liquid interface to mimic human skin such that the fibroblasts grow in culture but the keratinocytes don't. Immortalised human keratinocyte cell lines can be modified using retroviral shRNA (RNAi) to suppress PTCH1 and a BCC model is formed. By knocking out PTCH1 only, for example NEB1-shPTCH1, localised colony formation is seen similar to a nodBCC phenotype, but no stromal invasion is seen. Thus, it would be good to knock out or over express the mBCC genes too, such as knockout of HECTD4 within the keratinocyte, to see if it transforms it into a locally invasive mBCC phenotype. If a mBCC phenotype is obtained, therapeutic agents such as those described in the potential treatment targets could be added to the model to see if it modifies behaviour.

6.2.2 Deep sequencing driver gene analysis for SGC

Candidate driver genes were ascertained in a small population of 5 SGC, however, further analysis using deep sequencing on a larger population would help to refute or support these group of genes. Nineteen candidate genes were picked based on the Intogen and MutSigCV significance and biological relevance (see Table 6.1). Fifteen FFPE sample

have been identified for DNA extraction and will be sent for targeted, deep sequencing at x1000 depth.

Gene	Chromosome	Start position	End position	Length/bp
IL12RB2	chr1	67773046	67862583	89538
RPTN	chr1	152126070	152131704	5635
DNAH14	chr1	225117355	225586996	469642
IBSP	chr4	88720701	88733601	12901
GPX3	chr5	150399998	150408554	8557
SYNE1	chr6	152442818	152958534	515717
GPR31	chr6	167570359	167571319	961
REPIN1	chr7	150065878	150071133	5256
DLX6	chr7	96635289	96640352	5064
DNAH11	chr7	21582832	21941186	358355
CUBN	chr10	16865964	17171816	305853
ZNF37A	chr10	38383263	38412278	29016
FRA10AC1	chr10	95427639	95462329	34691
RB1	chr13	48877882	49056026	178145
GNG2	chr14	52327021	52436518	109498
C16orf87	chr16	46835958	46865074	29117
GJD3	chr17	38516904	38520945	4042
TP53	chr17	7571719	7590868	19150
PRPF31	chr19	54618789	54635150	16362

Table 6.1 Deep sequencing of 19 potential SGC driver genes. Selected driver genes for SGC deep sequencing on 15 samples. Probes are designed against the human reference genome 19 University of California Santa Cruz (UCSC) and National centre of biotechnology information (NCBI) 37.7.

6.2.3 Cancel panel analysis and personalised medicine

Driver genes from both our cohort and the Bonilla et al cohort could be further tested in a prospective study. At present, archival tissue is often used and it is still not clear what these changes mean to the patient in terms of prognosis or need for follow-up. By carrying out a cancer panel test encompassing the potential drivers on tissue samples in a prospective study, better data can be captured on its relevant to histology, recurrence risk and prognosis. Due to the low incidence of SGC, it would require an unfeasible length of time to collect enough patients. BCC is so frequent that it would be feasible to carry out a prospective study and depending on the power required, enough patients could be collected over the course of a 1 year.

Gene panels are becoming more common place in the NHS, but are mainly used to detect germline mutations for specific conditions such as congenital cataract or hearing loss.

Currently on the NHS, as an example, a blood sample (1-3ml) can be taken from a child with bilateral cataract, sent to a lab in Leeds where accredited DNA extraction, library preparation and Illumina Hiseq sequencing is performed on 114 known cataract causative genes. Bioinformatics is carried out and potentially pathogenic variants confirmed using conventional Sanger sequencing. A medical report is made of the clinical interpretation of the known variants identified. A technical report is made of the coverage along with a list of non-validated variants of unknown clinical significance. The whole process costs the NHS £1100 and has a fairly long turnaround time of 80 working days. This turnaround time has implications in cancer where adjunctive treatments such as chemotherapy may need to be started immediately after excision or even as a neoadjuvant therapy.

Cancer panels really started when the BRCA1 and BRCA2 were being offered to women who were at high risk developing breast or ovarian cancer. Since around 2013, there has been an expansion to test for other cancer related genes. Breast cancer is at the forefront of cancer panels and there are over 10 different companies providing a gene panel service, but the fact that we can test for them does not mean that the are clinically relevant for that individual, and in some cases could induce more unnecessary anxiety.(Easton et al., 2015) Cancer Research UK have recently approved the use of a custom panel containing 80 genes where the clinician can choose 19 genes pertaining to the patient's cancer for a fixed price of £860 and 16 week turnaround. This was developed after a large epidemiology project, the collaborative oncological gene-environment study (COGS) looked at 200,000 patients genome, half with cancer and the other half without and discovered 80 genes that conferred an increased risk of developing cancer.(Michailidou et al., 2013, Garcia-Closas et al., 2013, Bojesen et al., 2013, Pharoah et al., 2013, Eeles et al., 2013) This is still on a blood sample, hence will be looking at germline DNA and the risk of developing certain cancers or having a relapse. However, there are caveats with these gene panel analyses. Just over 100 patients have been analysed using the 80-gene panel and the rate of unknown variant detection currently stands at 25% (personal communication with Professor Eeles, Royal Marsden Hospital). Therefore, a quarter of the patients will be given an unnerving result at the end of the test and it is then not clear what to do next. Having such a high potentially false positive result may not lend itself to becoming a screening test in the near future.

The work carried out in this PhD focuses on the mutations found in the cancer rather than the germline, and as such a cancer panel relating to these genes would have to be tested on the biopsy sample using the surplus tissue not required for histological sampling. Fortunately, this technique has been perfected in our lab using laser capture microdissection and low input DNA extraction methods for WES. As mentioned earlier, the current lag time to detection of the variants from the biopsy and definitive excision / adjunctive treatment is an issue, however, it is hoped that with newer technology such as the minion, results can be obtained in less time. Personalised genetic medicine is already occurring in cancer, for example in melanoma, BRAF V600E or V600K positive patients, are eligible for MEK inhibitors (Trametinib [Mekinist]) or BRAF protein inhibitors (Vemurafenib [Zelboraf] and dabrafenib [Tafinlar]) under National Institute for Health and Care Excellence (NICE) guidelines. Having a personalised approach to cancer care and tailoring treatment according to the types of mutations or chromosomal instability will be the future of tumour therapy as more of the potential treatment targets mentioned in the previous chapters become more widely available.

Then there are the general concerns with genetic testing as we start to use these tools more often and screen more of the genome: who has access to this data? Can insurance companies ask for the details? Can family members request the results, especially if a germline test, as it may pertain to them too? Is the DNA stored indefinitely for future testing as technologies improve? Can companies or government agencies request access to this DNA for testing, for example the police in a criminal enquiry? With the advent of CRISP-Cas9 with its ability to modify sections of DNA with incredible accuracy, it is only a matter of time before we have to face these huge moral dilemmas of gene editing. Perhaps with cancer it is an easier answer, especially removal of an inherited cancer gene, so that the subsequent progeny will never be affected. Lets hope we, the international community, can all move forward together and beat cancer sooner.

Post scriptum

Cancer Research UK is the world's leading charity in beating cancer through research and advertises its Herculean task 'of bringing forward the day when all cancers are cured'. And what a day that would be, to have Cancer crushed beneath the foot of Hercules!

During my 18-month viva, I casually commented that cancer took over the cellular machinery for its own advantage and I was quickly shot down by the examiner who said that cancer cannot take over the body or its machinery. I reflected on this later on and I was still not happy with either viewpoint. It made me contemplate, what actually is cancer? As a scientist, I can talk about mutational advantage, survival of the fitness, consumption of the body's resources, destruction of normal tissue so it cannot function etc. However, as a doctor who cannot ignore the art of medicine, the patient who is diagnosed with cancer (the patient with cancer rather than the cancer patient) will never be the same again. Benign, indolent, aggressive, remission or terminal; yes they are very different in terms of outcome, but once the label is given, it cannot be taken away and is forever set within the mind. Of course, watching my Dad with terminal prostate cancer it may seem obvious. So going back to my flippant comment: does cancer take over the body? Yes, permanently. Either within the subconscious, preconscious or conscious mind. A good friend told me that doctors are the gatekeepers between land and the great abysses (of despair). Thus, a gatekeeper I shall be.

APPENDIX ONE

Individual gene level comparison of subtypes for selected pathways

	MORPHOEIC			NODULAR		
Gene	**Exp**	**LC**	**P-val**	**Exp**	**LC**	**P-val**
PTCH1	5.7	1.2	0.009	5.7	1.8	0.002
SMO	4.5	1.2	0.01	4.5	0.2	0.63
Gli1	4.9	1.5	0.004	4.9	1.8	0.006
Gli2	5.1	1.1	0.01	5.1	0.9	0.09
STK36	4.5	-0.1	0.86	4.5	0.1	0.778
LRP2	3.1	1.0	0.03	3.1	1.7	0.014
CREBBP	7.1	0.3	0.33	7.1	0.2	0.52
HHIP	3.6	2.3	0.001	3.7	1.4	0.021
SIN3A	6.3	0.2	0.69	6.3	0.2	0.6
BMP4	4.0	0.1	0.94	4.0	-0.4	0.64
CSNK1A1	5.6	-0.1	0.88	5.6	-0.3	0.44
CSNK1E	6.0	0.1	0.92	6.0	0.1	0.838
PRKACB	4.9	-0.5	0.22	4.9	-0.1	0.93
PRKX	6.8	0.9	0.03	6.8	0.5	0.13
WNT2	0.4	0.6	0.60	0.35	1.8	0.145

Table A1.1 Hedgehog pathway expression in mBCC and nodBCC. Mutational burden in 10 mBCC and 10 nodBCC tumours with respect to the Hedgehog signalling pathway (hsa04340). Exp=expression, P-val=P-valu

	MORPHOEIC			NODULAR		
Name	**Exp**	**LC**	**P-val**	**Exp**	**LC**	**P-val**
SLIT2	5.0	-0.6	0.22	5.0	0.5	0.34
EPHA3	1.8	-0.5	0.65	1.8	-0.4	0.72
EPHB1	0.9	0.3	0.80	0.9	-0.2	0.87
DCC	0.3	3.8	<0.01	0.3	0.7	0.41
LRRC4C	1.2	-2.1	0.15	1.2	0.1	0.96
ROBO2	3.1	-0.4	0.70	3.1	-0.1	0.98
SLIT3	6.4	0.5	0.15	6.4	0.4	0.36
UNC5C	-1.0	-0.3	0.85	-1.0	1.0	0.49
EPHA7	-0.7	-1.8	0.23	-0.8	0.8	0.56
FES	2.8	-1.0	0.25	2.8	-1.2	0.24
NGEF	-7.3	1.7	0.18	-0.7	-2.9	0.05
NFAT5	5.8	-0.2	0.53	5.8	0.1	0.72
NFATC2	5.5	-1.0	0.02	4.6	-0.1	0.89
PAK6	3.3	0.5	0.46	3.3	-0.3	0.72
PLXNC1	4.8	-1.3	0.03	4.8	-0.1	0.89
SEMA3E	1.9	0.1	0.97	3.0	0.1	0.97
SEMA4D	5.0	-0.6	0.21	5.0	-1.2	0.07
SEMA6C	0.7	0.1	0.92	0.7	1.0	0.99
ROCK1	7.3	-0.2	0.64	7.3	0.1	0.95

Table A1.2 Axonal guidance pathway expression in mBCC and nodBCC. Mutational burden in 10 mBCC and 10 nodBCC tumours with respect to the Hedgehog signalling pathway (hsa04340). Exp=expression, P-val=P-value

	MORPHOEIC			NODULAR		
Name	**Exp**	**LC**	**P-v**	**Exp**	**LC**	**P-v**
CREBBP	7.1	0.2	0.5	7.1	0.3	0.3
NFATC2	4.6	-0.1	0.9	5.5	-1.0	0.02
NFAT5	5.8	0.2	0.7	5.8	-0.2	0.5
RHOA	9.3	-0.1	0.7	9.3	-0.1	0.9
TBL1X	5.8	0.1	0.7	5.8	-0.8	0.1
PPP3R1	4.1	-0.1	0.9	4.1	0.2	0.8
MAPK9	5.0	0.4	0.4	4.9	0.1	0.8
LRP6	5.9	0.9	0.04	5.9	0.6	0.2
PRKACA	5.0	0.4	0.6	5.0	-0.1	0.8
EP300	7.0	0.2	0.7	3.1	-0.1	0.9
CHD8	7.1	-0.2	0.6	7.1	0.3	0.4
NFATC4	3.8	0.8	0.2	4.7	-0.6	0.3
CSNK2A1	6.9	-0.3	0.4	6.9	0.3	0.4
PORCN	2.3	-2.2	0.02	2.3	-0.5	0.7
AXIN1	3.0	-0.8	0.4	3.0	-0.1	0.9
FZD3	5.0	0.04	0.9	5.0	0.7	0.2
SFRP4	2.9	1.1	2.2	2.9	-1.2	0.3
DVL1	1.5	-1.3	0.3	1.5	0.3	0.7
MAPK8	4.6	0.1	0.8	1.5	0.3	0.7
WNT5B	2.5	-1.5	0.2	2.5	-1.2	0.2
FZD10	2.8	0.3	0.7	2.8	0.9	2.3
PPP2R5D	4.5	0.9	0.2	4.5	0.3	0.6
CSNK1A1	5.6	-0.3	0.5	5.6	-0.1	0.9
ROCK2	7.0	0.3	0.4	7.0	0.1	0.9
APC	6.4	0.5	0.2	6.4	-0.3	0.4
WNT10A	1.4	-0.1	0.9	1.4	-1.4	0.3
PPP2R1B	4.9	0.4	0.3	4.9	0.3	0.6
PPP3CA	7.2	0.2	0.5	7.2	-0.7	0.06
TP53	7.0	0.3	0.4	7.0	0.1	0.8
SMAD4	6.4	-0.2	0.6	6.4	0.1	0.8
TCF7L2	6.2	0.7	0.07	6.2	-0.1	0.9
PLCB3	2.1	-0.6	0.6	2.1	0.1	0.9
CTBP1	6.9	0.9	0.03	6.9	0.7	0.07
DVL3	5.9	0.4	0.3	5.9	-0.1	0.9
PRICKLE2	0.1	0.1	0.6	0.1	-0.9	0.4
SOX17	0.2	-2.2	0.1	-0.2	-1.5	0.3
PRKCB	2.6	-1.6	0.1	2.6	-1.6	0.1

Table A1.3 Wnt signalling pathway expression in mBCC and nodBCC. Mutational burden in 10 mBCC and 10 nodBCC tumours with respect to the Hedgehog signalling pathway (hsa04340). Exp=expression, P-val=P-value

	MORPHOEIC			NODULAR		
Name	**Exp**	**LC**	**P-val**	**Exp**	**LC**	**P-val**
INPP5D	3.8	-0.5	0.5	3.8	-1.4	0.03
FYN	6.5	0.3	0.5	6.5	-0.4	0.4
GAB2	4.0	-0.7	0.4	4.0	-0.7	0.4
MAP2K3	4.7	-0.6	0.3	4.7	-0.2	0.7
MAPK9	5.0	0.4	0.4	4.9	0.1	0.8
SOS2	5.5	0.4	0.3	5.6	-0.1	0.9
PIK3CG	0.7	-2.8	0.04	0.7	-2.4	0.05
MAPK8	4.8	0.1	0.8	4.9	-0.3	0.6
MAPK14	6.1	-0.4	0.3	6.1	-0.3	0.4
AKT3	4.3	-0.8	0.1	4.3	-0.5	0.4
VAV1	1.6	-2.5	0.09	1.6	-0.8	0.5
PIK3R1	7.8	0.7	0.07	7.8	0.7	0.08
MS4A2	0.9	1.7	0.3	0.9	-0.6	0.6
FCER1G	4.6	-1.5	0.04	4.6	-0.5	0.3
PRKCB	2.6	-1.6	0.1	2.6	-1.6	0.1
RAC3	1.2	1.2	0.4	1.2	1.7	0.2
PIK3CD	3.4	-1.0	0.3	3.4	-1.5	0.04
HRAS	3.9	-0.7	0.3	3.9	-0.4	0.5
PLA2G6	0.4	-0.3	0.8	0.4	1.1	0.4
PLCG2	4.1	0.1	0.9	4.1	-1.3	0.04

Table A1.4 Fc epsilon RI signalling expression in mBCC and nodBCC. Mutational burden in 10 mBCC tumours with respect to the Hedgehog signalling pathway (hsa04340). Exp=expression, P-val=P-value

REFERENCES

ABDELSALAM, M., EL SISSY, A., SAMRA, M. A., IBRAHIM, S., EL MARKABY, D. & GADALLAH, F. 2008. The impact of trisomy 12, retinoblastoma gene and P53 in prognosis of B-cell chronic lymphocytic leukemia. *Hematology,* 13**,** 147-53.

ADAMS, E. J. 2013. Diverse antigen presentation by the Group 1 CD1 molecule, CD1c. *Mol Immunol,* 55**,** 182-5.

ADJEI, A. A. 2001. Gemcitabine and pemetrexed disodium combinations in vitro and in vivo. *Lung Cancer,* 34 Suppl 4**,** S103-5.

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nat Methods.* United States.

AGNIHOTRI, N., KUMAR, S. & MEHTA, K. 2013. Tissue transglutaminase as a central mediator in inflammation-induced progression of breast cancer. *Breast Cancer Res,* 15**,** 202.

AHN, J. Y., LEE, J. S., MIN, H. Y. & LEE, H. Y. 2015. Acquired resistance to 5-fluorouracil via HSP90/Src-mediated increase in thymidylate synthase expression in colon cancer. *Oncotarget*.

AIROLDI, I., DI CARLO, E., COCCO, C., SORRENTINO, C., FAIS, F., CILLI, M., D'ANTUONO, T., COLOMBO, M. P. & PISTOIA, V. 2005. Lack of Il12rb2 signaling predisposes to spontaneous autoimmunity and malignancy. *Blood,* 106**,** 3846-53.

AKAGI, T., ITO, T., KATO, M., JIN, Z., CHENG, Y., KAN, T., YAMAMOTO, G., OLARU, A., KAWAMATA, N., BOULT, J., SOUKIASIAN, H. J., MILLER, C. W., OGAWA, S., MELTZER, S. J. & KOEFFLER, H. P. 2009. Chromosomal abnormalities and novel disease-related regions in progression from Barrett's esophagus to esophageal adenocarcinoma. *Int J Cancer,* 125**,** 2349-59.

AL-HAJJ, M., WICHA, M. S., BENITO-HERNANDEZ, A., MORRISON, S. J. & CLARKE, M. F. 2003. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A,* 100**,** 3983-8.

AL-KHALAF, H. H., HENDRAYANI, S. F. & ABOUSSEKHRA, A. 2012. ATR controls the p21(WAF1/Cip1) protein up-regulation and apoptosis in response to low UV fluences. *Mol Carcinog,* 51**,** 930-8.

AL-TASSAN, N. A., WHIFFIN, N., HOSKING, F. J., PALLES, C., FARRINGTON, S. M., DOBBINS, S. E., HARRIS, R., GORMAN, M., TENESA, A., MEYER, B. F., WAKIL, S. M., KINNERSLEY, B., CAMPBELL, H., MARTIN, L., SMITH, C. G., IDZIASZCZYK, S., BARCLAY, E., MAUGHAN, T. S., KAPLAN, R., KERR, R., KERR, D., BUCHANAN, D. D., WIN, A. K., HOPPER, J., JENKINS, M., LINDOR, N. M., NEWCOMB, P. A., GALLINGER, S., CONTI, D., SCHUMACHER, F., CASEY, G., DUNLOP, M. G., TOMLINSON, I. P., CHEADLE, J. P. & HOULSTON, R. S. 2015. A new GWAS and meta-analysis with 1000Genomes imputation identifies novel risk variants for colorectal cancer. *Sci Rep,* 5**,** 10442.

ALEXAKI, V. I., JAVELAUD, D., VAN KEMPEN, L. C., MOHAMMAD, K. S., DENNLER, S., LUCIANI, F., HOEK, K. S., JUAREZ, P., GOYDOS, J. S.,

FOURNIER, P. J., SIBON, C., BERTOLOTTO, C., VERRECCHIA, F., SAULE, S., DELMAS, V., BALLOTTI, R., LARUE, L., SAIAG, P., GUISE, T. A. & MAUVIEL, A. 2010. GLI2-mediated melanoma invasion and metastasis. *J Natl Cancer Inst.* United States.

ALEXANDROV, L. B., NIK-ZAINAL, S., WEDGE, D. C., APARICIO, S. A., BEHJATI, S., BIANKIN, A. V., BIGNELL, G. R., BOLLI, N., BORG, A., BORRESEN-DALE, A. L., BOYAULT, S., BURKHARDT, B., BUTLER, A. P., CALDAS, C., DAVIES, H. R., DESMEDT, C., EILS, R., EYFJORD, J. E., FOEKENS, J. A., GREAVES, M., HOSODA, F., HUTTER, B., ILICIC, T., IMBEAUD, S., IMIELINSKI, M., JAGER, N., JONES, D. T., JONES, D., KNAPPSKOG, S., KOOL, M., LAKHANI, S. R., LOPEZ-OTIN, C., MARTIN, S., MUNSHI, N. C., NAKAMURA, H., NORTHCOTT, P. A., PAJIC, M., PAPAEMMANUIL, E., PARADISO, A., PEARSON, J. V., PUENTE, X. S., RAINE, K., RAMAKRISHNA, M., RICHARDSON, A. L., RICHTER, J., ROSENSTIEL, P., SCHLESNER, M., SCHUMACHER, T. N., SPAN, P. N., TEAGUE, J. W., TOTOKI, Y., TUTT, A. N., VALDES-MAS, R., VAN BUUREN, M. M., VAN 'T VEER, L., VINCENT-SALOMON, A., WADDELL, N., YATES, L. R., ZUCMAN-ROSSI, J., FUTREAL, P. A., MCDERMOTT, U., LICHTER, P., MEYERSON, M., GRIMMOND, S. M., SIEBERT, R., CAMPO, E., SHIBATA, T., PFISTER, S. M., CAMPBELL, P. J. & STRATTON, M. R. 2013. Signatures of mutational processes in human cancer. *Nature.* England.

ALI-RIDHA, A. B., S; JIANG, K; MILMAN, T; BURNS,B; BLANCO, P; FARMER, J 2014. Immunohistochemical Analysis of Sebaceous Cell Carcinoma in Comparison to Both Basal Cell Carcinoma and Squamous Cell Carcinoma. *Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO).* Orlando, USA: ARVO.

ALLY, M. S., AASI, S., WYSONG, A., TENG, C., ANDERSON, E., BAILEY-HEALY, I., ORO, A., KIM, J., CHANG, A. L. & TANG, J. Y. 2014. An investigator-initiated open-label clinical trial of vismodegib as a neoadjuvant to surgery for high-risk basal cell carcinoma. *J Am Acad Dermatol,* 71**,** 904-911 e1.

ALVAREZ-MEDINA, R., CAYUSO, J., OKUBO, T., TAKADA, S. & MARTI, E. 2008. Wnt canonical pathway restricts graded Shh/Gli patterning activity through the regulation of Gli3 expression. *Development,* 135**,** 237-47.

AMATSCHEK, S., LUCAS, R., EGER, A., PFLUEGER, M., HUNDSBERGER, H., KNOLL, C., GROSSE-KRACHT, S., SCHUETT, W., KOSZIK, F., MAURER, D. & WIESNER, C. 2011. CXCL9 induces chemotaxis, chemorepulsion and endothelial barrier disruption through CXCR3-mediated activation of melanoma cells. *Br J Cancer,* 104**,** 469-79.

AMINOFF, M., CARTER, J. E., CHADWICK, R. B., JOHNSON, C., GRASBECK, R., ABDELAAL, M. A., BROCH, H., JENNER, L. B., VERROUST, P. J., MOESTRUP, S. K., DE LA CHAPELLE, A. & KRAHE, R. 1999. Mutations in CUBN, encoding the intrinsic factor-vitamin B12 receptor, cubilin, cause hereditary megaloblastic anaemia 1. *Nat Genet,* 21**,** 309-13.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-9.

ANDREE, K. C., BARRADAS, A. M., NGUYEN, A. T., MENTINK, A., STOJANOVIC, I., BAGGERMAN, J., VAN DALUM, J., VAN RIJN, C. J. & TERSTAPPEN, L. W. 2016. Capture of Tumor Cells on Anti-EpCAM-

Functionalized Poly(acrylic acid)-Coated Surfaces. *ACS Appl Mater Interfaces,* 8**,** 14349-56.

ANDREWS, S. *FastQC A quality control tool for high throughput sequence data* [Online]. Babraham Bioinformatics. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed].

ANDREWS, S. J. & ROTHNAGEL, J. A. 2014. Emerging evidence for functional peptides encoded by short open reading frames. *Nat Rev Genet,* 15**,** 193-204.

ANTON, I. M., DE LA FUENTE, M. A., SIMS, T. N., FREEMAN, S., RAMESH, N., HARTWIG, J. H., DUSTIN, M. L. & GEHA, R. S. 2002. WIP deficiency reveals a differential role for WIP and the actin cytoskeleton in T and B cell activation. *Immunity,* 16**,** 193-204.

APORTA, A., CATALAN, E., GALAN-MALO, P., RAMIREZ-LABRADA, A., PEREZ, M., AZACETA, G., PALOMERA, L., NAVAL, J., MARZO, I., PARDO, J. & ANEL, A. 2014. Granulysin induces apoptotic cell death and cleavage of the autophagy regulator Atg5 in human hematological tumors. *Biochem Pharmacol,* 87**,** 410-23.

ARAI, E., GOTOH, M., TIAN, Y., SAKAMOTO, H., ONO, M., MATSUDA, A., TAKAHASHI, Y., MIYATA, S., TOTSUKA, H., CHIKU, S., KOMIYAMA, M., FUJIMOTO, H., MATSUMOTO, K., YAMADA, T., YOSHIDA, T. & KANAI, Y. 2015. Alterations of the spindle checkpoint pathway in clinicopathologically aggressive CpG island methylator phenotype clear cell renal cell carcinomas. *Int J Cancer,* 137**,** 2589-606.

ARTHUR, W. T. & BURRIDGE, K. 2001. RhoA inactivation by p190RhoGAP regulates cell spreading and migration by promoting membrane protrusion and polarity. *Mol Biol Cell,* 12**,** 2711-20.

ARUMUGAM, T., RAMACHANDRAN, V., SUN, D., PENG, Z., PAL, A., MAXWELL, D. S., BORNMANN, W. G. & LOGSDON, C. D. 2013. Designing and developing S100P inhibitor 5-methyl cromolyn for pancreatic cancer therapy. *Mol Cancer Ther,* 12**,** 654-62.

ASP, P., WIHLBORG, M., KARLEN, M. & FARRANTS, A. K. 2002. Expression of BRG1, a human SWI/SNF component, affects the organisation of actin filaments through the RhoA signalling pathway. *J Cell Sci,* 115**,** 2735-46.

AYAZ, L., GORUR, A., YAROGLU, H. Y., OZCAN, C. & TAMER, L. 2013. Differential expression of microRNAs in plasma of patients with laryngeal squamous cell carcinoma: potential early-detection markers for laryngeal squamous cell carcinoma. *J Cancer Res Clin Oncol,* 139**,** 1499-506.

BABA, M., HONG, S. B., SHARMA, N., WARREN, M. B., NICKERSON, M. L., IWAMATSU, A., ESPOSITO, D., GILLETTE, W. K., HOPKINS, R. F., 3RD, HARTLEY, J. L., FURIHATA, M., OISHI, S., ZHEN, W., BURKE, T. R., JR., LINEHAN, W. M., SCHMIDT, L. S. & ZBAR, B. 2006. Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. *Proc Natl Acad Sci U S A,* 103**,** 15552-7.

BAFICO, A., GAZIT, A., PRAMILA, T., FINCH, P. W., YANIV, A. & AARONSON, S. A. 1999. Interaction of frizzled related protein (FRP) with Wnt ligands and the frizzled receptor suggests alternative mechanisms for FRP inhibition of Wnt signaling. *J Biol Chem,* 274**,** 16180-7.

BAHAT, A., KEDMI, R., GAZIT, K., RICHARDO-LAX, I., AINBINDER, E. & DIKSTEIN, R. 2013. TAF4b and TAF4 differentially regulate mouse embryonic stem cells maintenance and proliferation. *Genes Cells,* 18**,** 225-37.

BAI, Z. L., FENG, Y. G., TAN, S. S., WANG, X. Y., XIAO, S. X., WANG, H., JIA, H. Q., WU, J. W., HE, D. L. & KANG, R. H. 2008. Mutations of KRT6A are more frequent than those of KRT16 in pachyonychia congenita type 1: report of a novel and a recently reported mutation in two unrelated Chinese families. *Br J Dermatol.* England.

BALL JR, A. R. & YOKOMORI, K. 2001. The structural maintenance of chromosomes (SMC) family of proteins in mammals. *Chromosome Res,* 9**,** 85-96.

BAMFORD, S., DAWSON, E., FORBES, S., CLEMENTS, J., PETTETT, R., DOGAN, A., FLANAGAN, A., TEAGUE, J., FUTREAL, P. A., STRATTON, M. R. & WOOSTER, R. 2004. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer,* 91**,** 355-8.

BANON-RODRIGUEZ, I., SAEZ DE GUINOA, J., BERNARDINI, A., RAGAZZINI, C., FERNANDEZ, E., CARRASCO, Y. R., JONES, G. E., WANDOSELL, F. & ANTON, I. M. 2013. WIP regulates persistence of cell migration and ruffle formation in both mesenchymal and amoeboid modes of motility. *PLoS One,* 8**,** e70364.

BARBIE, V. & LEFRANC, M. P. 1998. The human immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments. *Exp Clin Immunogenet,* 15**,** 171-83.

BARQUILLA, A. & PASQUALE, E. B. 2015. Eph receptors and ephrins: therapeutic opportunities. *Annu Rev Pharmacol Toxicol,* 55**,** 465-87.

BARRETT, C. W., NING, W., CHEN, X., SMITH, J. J., WASHINGTON, M. K., HILL, K. E., COBURN, L. A., PEEK, R. M., CHATURVEDI, R., WILSON, K. T., BURK, R. F. & WILLIAMS, C. S. 2013. Tumor suppressor function of the plasma glutathione peroxidase gpx3 in colitis-associated carcinoma. *Cancer Res,* 73**,** 1245-55.

BARRON, E. V., ROMAN-BASSAURE, E., SANCHEZ-SANDOVAL, A. L., ESPINOSA, A. M., GUARDADO-ESTRADA, M., MEDINA, I., JUAREZ, E., ALFARO, A., BERMUDEZ, M., ZAMORA, R., GARCIA-RUIZ, C., GOMORA, J. C., KOFMAN, S., PEREZ-ARMENDARIZ, E. M. & BERUMEN, J. 2015. CDKN3 mRNA as a Biomarker for Survival and Therapeutic Target in Cervical Cancer. *PLoS One,* 10**,** e0137397.

BARTHOLOMEUSZ, G., WU, Y., ALI SEYED, M., XIA, W., KWONG, K. Y., HORTOBAGYI, G. & HUNG, M. C. 2006. Nuclear translocation of the pro-apoptotic Bcl-2 family member Bok induces apoptosis. *Mol Carcinog,* 45**,** 73-83.

BASS-ZUBEK, A. E., GODSEL, L. M., DELMAR, M. & GREEN, K. J. 2009. Plakophilins: multifunctional scaffolds for adhesion and signaling. *Curr Opin Cell Biol,* 21**,** 708-16.

BAYLIN, S. B. & JONES, P. A. 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer,* 11**,** 726-34.

BECKER, E. H., M; HOLZ, F; FISCHER, HP; LOEFFLER, K 2014. Sebaceous gland carcinoma of the ocular adnexa – variability in histological and immunhistochemical appearance. *The Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO)*. Orlando, USA: ARVO.

BELLEFROID, E. J., MARINE, J. C., RIED, T., LECOCQ, P. J., RIVIERE, M., AMEMIYA, C., PONCELET, D. A., COULIE, P. G., DE JONG, P., SZPIRER, C. & ET AL. 1993. Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *Embo j,* 12**,** 1363-74.

BELLESSORT, B., LE CARDINAL, M., BACHELOT, A., NARBOUX-NEME, N., GARAGNANI, P., PIRAZZINI, C., BARBIERI, O., MASTRACCI, L., JONCHERE, V., DUVERNOIS-BERTHET, E., FONTAINE, A., ALFAMA, G. & LEVI, G. 2016. Dlx5 and Dlx6 control uterine adenogenesis during post-natal maturation: possible consequences for endometriosis. *Hum Mol Genet,* 25**,** 97-108.

BENVENUTO, M., MASUELLI, L., DE SMAELE, E., FANTINI, M., MATTERA, R., CUCCHI, D., BONANNO, E., DI STEFANO, E., FRAJESE, G. V., ORLANDI, A., SCREPANTI, I., GULINO, A., MODESTI, A. & BEI, R. 2016. In vitro and in vivo inhibition of breast cancer cell growth by targeting the Hedgehog/GLI pathway with SMO (GDC-0449) or GLI (GANT-61) inhibitors. *Oncotarget,* 7**,** 9250-70.

BERG, D. & OTLEY, C. C. 2002. Skin cancer in organ transplant recipients: Epidemiology, pathogenesis, and management. *J Am Acad Dermatol,* 47**,** 1-17; quiz 18-20.

BERGETHON, K., SHAW, A. T., OU, S. H., KATAYAMA, R., LOVLY, C. M., MCDONALD, N. T., MASSION, P. P., SIWAK-TAPP, C., GONZALEZ, A., FANG, R., MARK, E. J., BATTEN, J. M., CHEN, H., WILNER, K. D., KWAK, E. L., CLARK, J. W., CARBONE, D. P., JI, H., ENGELMAN, J. A., MINO-KENUDSON, M., PAO, W. & IAFRATE, A. J. 2012. ROS1 rearrangements define a unique molecular class of lung cancers. *J Clin Oncol,* 30**,** 863-70.

BERKOVA, Z., TAO, R. H. & SAMANIEGO, F. 2010. Milatuzumab - a promising new immunotherapeutic agent. *Expert Opin Investig Drugs,* 19**,** 141-9.

BERMAN, D. M., KARHADKAR, S. S., MAITRA, A., MONTES DE OCA, R., GERSTENBLITH, M. R., BRIGGS, K., PARKER, A. R., SHIMADA, Y., ESHLEMAN, J. R., WATKINS, D. N. & BEACHY, P. A. 2003. Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours. *Nature,* 425**,** 846-51.

BERNTSSON, J., LUNDGREN, S., NODIN, B., UHLEN, M., GABER, A. & JIRSTROM, K. 2014. Expression and prognostic significance of the polymeric immunoglobulin receptor in epithelial ovarian cancer. *J Ovarian Res,* 7**,** 26.

BHARDWAJ, M., SEN, S., SHARMA, A., KASHYAP, S., CHOSDOL, K., PUSHKER, N., BAJAJ, M. S. & BAKHSHI, S. 2015. ZEB2/SIP1 as novel prognostic indicator in eyelid sebaceous gland carcinoma. *Hum Pathol,* 46**,** 1437-42.

BHARTIYA 2015. lncRNome | long noncoding RNA Knowledgebase.

BIRD, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev,* 16**,** 6-21.

BLOOR, B. K., SEDDON, S. V. & MORGAN, P. R. 2001. Gene expression of differentiation-specific keratins in oral epithelial dysplasia and squamous cell carcinoma. *Oral Oncol,* 37**,** 251-61.

BOCK, V. L., LYONS, J. G., HUANG, X. X., JONES, A. M., MCDONALD, L. A., SCOLYER, R. A., MOLONEY, F. J., BARNETSON, R. S. & HALLIDAY, G. M. 2011. BRM and BRG1 subunits of the SWI/SNF chromatin remodelling complex are downregulated upon progression of benign skin lesions into invasive tumours. *Br J Dermatol,* 164**,** 1221-7.

BOERNER, J. L., DEMORY, M. L., SILVA, C. & PARSONS, S. J. 2004. Phosphorylation of Y845 on the epidermal growth factor receptor mediates

binding to the mitochondrial protein cytochrome c oxidase subunit II. *Mol Cell Biol,* 24**,** 7059-71.

BOEVA, V., POPOVA, T., LIENARD, M., TOFFOLI, S., KAMAL, M., LE TOURNEAU, C., GENTIEN, D., SERVANT, N., GESTRAUD, P., RIO FRIO, T., HUPÉ, P., BARILLOT, E. & LAES, J. F. 2014. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinformatics.*

BOJESEN, S. E., POOLEY, K. A., JOHNATTY, S. E., BEESLEY, J., MICHAILIDOU, K., TYRER, J. P., EDWARDS, S. L., PICKETT, H. A., SHEN, H. C., SMART, C. E., HILLMAN, K. M., MAI, P. L., LAWRENSON, K., STUTZ, M. D., LU, Y., KAREVAN, R., WOODS, N., JOHNSTON, R. L., FRENCH, J. D., CHEN, X., WEISCHER, M., NIELSEN, S. F., MARANIAN, M. J., GHOUSSAINI, M., AHMED, S., BAYNES, C., BOLLA, M. K., WANG, Q., DENNIS, J., MCGUFFOG, L., BARROWDALE, D., LEE, A., HEALEY, S., LUSH, M., TESSIER, D. C., VINCENT, D., BACOT, F., VERGOTE, I., LAMBRECHTS, S., DESPIERRE, E., RISCH, H. A., GONZALEZ-NEIRA, A., ROSSING, M. A., PITA, G., DOHERTY, J. A., ALVAREZ, N., LARSON, M. C., FRIDLEY, B. L., SCHOOF, N., CHANG-CLAUDE, J., CICEK, M. S., PETO, J., KALLI, K. R., BROEKS, A., ARMASU, S. M., SCHMIDT, M. K., BRAAF, L. M., WINTERHOFF, B., NEVANLINNA, H., KONECNY, G. E., LAMBRECHTS, D., ROGMANN, L., GUENEL, P., TEOMAN, A., MILNE, R. L., GARCIA, J. J., COX, A., SHRIDHAR, V., BURWINKEL, B., MARME, F., HEIN, R., SAWYER, E. J., HAIMAN, C. A., WANG-GOHRKE, S., ANDRULIS, I. L., MOYSICH, K. B., HOPPER, J. L., ODUNSI, K., LINDBLOM, A., GILES, G. G., BRENNER, H., SIMARD, J., LURIE, G., FASCHING, P. A., CARNEY, M. E., RADICE, P., WILKENS, L. R., SWERDLOW, A., GOODMAN, M. T., BRAUCH, H., GARCIA-CLOSAS, M., HILLEMANNS, P., WINQVIST, R., DURST, M., DEVILEE, P., RUNNEBAUM, I., JAKUBOWSKA, A., LUBINSKI, J., MANNERMAA, A., BUTZOW, R., et al. 2013. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet,* 45**,** 371-84, 384e1-2.

BOLOGNA-MOLINA, R., MOSQUEDA-TAYLOR, A. & MOLINA-FRECHERO, N. 2015. Differential expression of glypican-1 in ameloblastoma variants. *Appl Immunohistochem Mol Morphol,* 23**,** 153-60.

BONILLA, X., PARMENTIER, L., KING, B., BEZRUKOV, F., KAYA, G., ZOETE, V., SEPLYARSKIY, V. B., SHARPE, H. J., MCKEE, T., LETOURNEAU, A., RIBAUX, P. G., POPADIN, K., BASSET-SEGUIN, N., CHAABENE, R. B., SANTONI, F. A., ANDRIANOVA, M. A., GUIPPONI, M., GARIERI, M., VERDAN, C., GROSDEMANGE, K., SUMARA, O., EILERS, M., AIFANTIS, I., MICHIELIN, O., SAUVAGE, F. J. D., ANTONARAKIS, S. E. & NIKOLAEV, S. I. 2016. Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nature Genetics*.

BONIUK, M. & ZIMMERMAN, L. E. 1968. Sebaceous carcinoma of the eyelid, eyebrow, caruncle, and orbit. *Trans Am Acad Ophthalmol Otolaryngol,* 72**,** 619-42.

BOONJARASPINYO, S., WU, Z., BOONMARS, T., KAEWKES, S., LOILOME, W., SITHITHAWORN, P., NAGANO, I., TAKAHASHI, Y., YONGVANIT, P. & BHUDHISAWASDI, V. 2012. Overexpression of PDGFA and its receptor

during carcinogenesis of Opisthorchis viverrini-associated cholangiocarcinoma. *Parasitol Int,* 61**,** 145-50.

BORLAK, J., LANGER, F., SPANEL, R., SCHONDORFER, G. & DITTRICH, C. 2016. Immune-mediated liver injury of the cancer therapeutic antibody catumaxomab targeting EpCAM, CD3 and Fcgamma receptors. *Oncotarget*.

BOT, B. M., ECKEL-PASSOW, J. E., LEGRAND, S. N., HILTON, T., CHEVILLE, J. C., IGEL, T. & PARKER, A. S. 2012. Expression of endothelin 2 and localized clear cell renal cell carcinoma. *Hum Pathol,* 43**,** 843-9.

BOUTET, A., COMAI, G. & SCHEDL, A. 2010. The WTX/AMER1 gene family: evolution, signature and function. *BMC Evol Biol,* 10**,** 280.

BOYD, A. W., BARTLETT, P. F. & LACKMANN, M. 2014. Therapeutic targeting of EPH receptors and their ligands. *Nat Rev Drug Discov,* 13**,** 39-62.

BRAUN, R., FINNEY, R., YAN, C., CHEN, Q. R., HU, Y., EDMONSON, M., MEERZAMAN, D. & BUETOW, K. 2013. Discovery Analysis of TCGA Data Reveals Association between Germline Genotype and Survival in Ovarian Cancer Patients. *PLoS One.*

BREUNINGER, H. & DIETZ, K. 1991. Prediction of subclinical tumor infiltration in basal cell carcinoma. *J Dermatol Surg Oncol,* 17**,** 574-8.

BRISCOE, J. & THEROND, P. P. 2013. The mechanisms of Hedgehog signalling and its roles in development and disease. *Nat Rev Mol Cell Biol,* 14**,** 416-29.

BRITZEN-LAURENT, N., LIPNIK, K., OCKER, M., NASCHBERGER, E., SCHELLERER, V. S., CRONER, R. S., VIETH, M., WALDNER, M., STEINBERG, P., HOHENADL, C. & STURZL, M. 2013. GBP-1 acts as a tumor suppressor in colorectal cancer cells. *Carcinogenesis,* 34**,** 153-62.

BUFALINO, A., CERVIGNE, N. K., DE OLIVEIRA, C. E., FONSECA, F. P., RODRIGUES, P. C., MACEDO, C. C., SOBRAL, L. M., MIGUEL, M. C., LOPES, M. A., PAES LEME, A. F., LAMBERT, D. W., SALO, T. A., KOWALSKI, L. P., GRANER, E. & COLETTA, R. D. 2015. Low miR-143/miR-145 Cluster Levels Induce Activin A Overexpression in Oral Squamous Cell Carcinomas, Which Contributes to Poor Prognosis. *PLoS One,* 10**,** e0136599.

BUITRAGO, W. & JOSEPH, A. K. 2008. Sebaceous carcinoma: the great masquerader: emgerging concepts in diagnosis and treatment. *Dermatol Ther,* 21**,** 459-66.

BULTMAN, S., GEBUHR, T., YEE, D., LA MANTIA, C., NICHOLSON, J., GILLIAM, A., RANDAZZO, F., METZGER, D., CHAMBON, P., CRABTREE, G. & MAGNUSON, T. 2000. A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol Cell,* 6**,** 1287-95.

BUNN, P. A., JR. & FRANKLIN, W. 2002. Epidermal growth factor receptor expression, signal pathway, and inhibitors in non-small cell lung cancer. *Semin Oncol,* 29**,** 38-44.

BURDELSKI, C., STRAUSS, C., TSOURLAKIS, M. C., KLUTH, M., HUBE-MAGG, C., MELLING, N., LEBOK, P., MINNER, S., KOOP, C., GRAEFEN, M., HEINZER, H., WITTMER, C., KRECH, T., SAUTER, G., WILCZAK, W., SIMON, R., SCHLOMM, T. & STEURER, S. 2015. Overexpression of thymidylate synthase (TYMS) is associated with aggressive tumor features and early PSA recurrence in prostate cancer. *Oncotarget,* 6**,** 8377-87.

BUSCHMANN, W. 2002. A reappraisal of cryosurgery for eyelid basal cell carcinomas. *Br J Ophthalmol,* 86**,** 453-7.

BYUN, S. Y., SHIN, Y. J., NAM, K. Y., HONG, S. P. & AHN, S. K. 2015. A novel highly potent and selective 11beta-hydroxysteroid dehydrogenase type 1 inhibitor, UI-1499. *Life Sci,* 120**,** 1-7.

CABALLERO, O. L. & CHEN, Y. T. 2009. Cancer/testis (CT) antigens: potential targets for immunotherapy. *Cancer Sci,* 100**,** 2014-21.

CAI, J., FANG, L., HUANG, Y., LI, R., YUAN, J., YANG, Y., ZHU, X., CHEN, B., WU, J. & LI, M. 2013. miR-205 targets PTEN and PHLPP2 to augment AKT signaling and drive malignant phenotypes in non-small cell lung cancer. *Cancer Res,* 73**,** 5402-15.

CAI, L., YUAN, W., ZHANG, Z., HE, L. & CHOU, K. C. 2016. In-depth comparison of somatic point mutation callers based on different tumor next-generation sequencing depth data. *Sci Rep,* 6**,** 36540.

CAI, Y., GEUTJES, E. J., DE LINT, K., ROEPMAN, P., BRUURS, L., YU, L. R., WANG, W., VAN BLIJSWIJK, J., MOHAMMAD, H., DE RINK, I., BERNARDS, R. & BAYLIN, S. B. 2014. The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. *Oncogene,* 33**,** 2157-68.

CANTILENA, S., PASTORINO, F., PEZZOLO, A., CHAYKA, O., PISTOIA, V., PONZONI, M. & SALA, A. 2011. Frizzled receptor 6 marks rare, highly tumourigenic stem-like cells in mouse and human neuroblastomas. *Oncotarget,* 2**,** 976-83.

CAO, W. J., WU, H. L., HE, B. S., ZHANG, Y. S. & ZHANG, Z. Y. 2013. Analysis of long non-coding RNA expression profiles in gastric cancer. *World J Gastroenterol,* 19**,** 3658-64.

CAPORALI, A. & EMANUELI, C. 2011. MicroRNA-503 and the extended microRNA-16 family in angiogenesis. *Trends Cardiovasc Med,* 21**,** 162-6.

CASTRO-SANCHEZ, L., SOTO-GUZMAN, A., GUADERRAMA-DIAZ, M., CORTES-REYNOSA, P. & SALAZAR, E. P. 2011. Role of DDR1 in the gelatinases secretion induced by native type IV collagen in MDA-MB-231 breast cancer cells. *Clin Exp Metastasis,* 28**,** 463-77.

CAVALLERO, S., SHEN, H., YI, C., LIEN, C. L., KUMAR, S. R. & SUCOV, H. M. 2015. CXCL12 Signaling Is Essential for Maturation of the Ventricular Coronary Endothelial Plexus and Establishment of Functional Coronary Circulation. *Dev Cell,* 33**,** 469-77.

CAVENEE, W. K., DRYJA, T. P., PHILLIPS, R. A., BENEDICT, W. F., GODBOUT, R., GALLIE, B. L., MURPHREE, A. L., STRONG, L. C. & WHITE, R. L. 1983. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature,* 305**,** 779-84.

CERUTTI, J. M., OLER, G., DELCELO, R., GERARDT, R., MICHALUART, P., JR., DE SOUZA, S. J., GALANTE, P. A., HUANG, P. & RIGGINS, G. J. 2011. PVALB, a new Hurthle adenoma diagnostic marker identified through gene expression. *J Clin Endocrinol Metab,* 96**,** E151-60.

CHAKRABORTY, C., DUTTA, S., MUKHERJEE, N., SAMADDER, S., ROYCHOWDHURY, A., ROY, A., MONDAL, R. K., BASU, P., ROYCHOUDHURY, S. & PANDA, C. K. 2015. Inactivation of PTCH1 is associated with the development of cervical carcinoma: clinical and prognostic implication. *Tumour Biol,* 36**,** 1143-54.

CHALIGNE, R. & HEARD, E. 2014. X-chromosome inactivation in development and cancer. *FEBS Lett,* 588**,** 2514-22.

CHANG, A. L. S. & ORO, A. E. 2012. Initial Assessment of Tumor Regrowth After Vismodegib in Advanced Basal Cell Carcinoma. *Arch Dermatol,* 148**,** 1324-5.

CHANG, X., HAN, J., PANG, L., ZHAO, Y., YANG, Y. & SHEN, Z. 2009. Increased PADI4 expression in blood and tissues of patients with malignant tumors. *BMC Cancer,* 9**,** 40.

CHANG, Y. Y., KUO, W. H., HUNG, J. H., LEE, C. Y., LEE, Y. H., CHANG, Y. C., LIN, W. C., SHEN, C. Y., HUANG, C. S., HSIEH, F. J., LAI, L. C., TSAI, M. H., CHANG, K. J. & CHUANG, E. Y. 2015. Deregulated microRNAs in triple-negative breast cancer revealed by deep sequencing. *Mol Cancer,* 14**,** 36.

CHAO, A. N., SHIELDS, C. L., KREMA, H. & SHIELDS, J. A. 2001. Outcome of patients with periocular sebaceous gland carcinoma with and without conjunctival intraepithelial invasion. *Ophthalmology,* 108**,** 1877-83.

CHAPNIK, E., RIVKIN, N., MILDNER, A., BECK, G., PASVOLSKY, R., METZL-RAZ, E., BIRGER, Y., AMIR, G., TIROSH, I., PORAT, Z., ISRAEL, L. L., LELLOUCHE, E., MICHAELI, S., LELLOUCHE, J. P., IZRAELI, S., JUNG, S. & HORNSTEIN, E. 2014. miR-142 orchestrates a network of actin cytoskeleton regulators during megakaryopoiesis. *Elife,* 3**,** e01964.

CHAUDHRY, M. A. 2013. Expression pattern of small nucleolar RNA host genes and long non-coding RNA in X-rays-treated lymphoblastoid cells. *Int J Mol Sci,* 14**,** 9099-110.

CHAVANAS, S., ADOUE, V., MECHIN, M. C., YING, S., DONG, S., DUPLAN, H., CHARVERON, M., TAKAHARA, H., SERRE, G. & SIMON, M. 2008. Long-range enhancer associated with chromatin looping allows AP-1 regulation of the peptidylarginine deiminase 3 gene in differentiated keratinocyte. *PLoS One,* 3**,** e3408.

CHEISHVILI, D., STEFANSKA, B., YI, C., CHEN LI, C., YU, P., ARAKELIAN, A., TANVIR, I., AHMED KHAN, H., RABBANI, S. & SZYF, M. 2015. A common promoter hypomethylation signature in invasive breast, liver and prostate cancer cell lines reveals novel targets involved in cancer invasiveness. *Oncotarget*.

CHEN, C., ZHOU, Z., LIU, R., LI, Y., AZMI, P. B. & SETH, A. K. 2008. The WW domain containing E3 ubiquitin protein ligase 1 upregulates ErbB2 and EGFR through RING finger protein 11. *Oncogene,* 27**,** 6845-55.

CHEN, Q. & ZHONG, T. 2015. The association of CXCR4 expression with clinicopathological significance and potential drug target in prostate cancer: a meta-analysis and literature review. *Drug Des Devel Ther,* 9**,** 5115-22.

CHEN, S. & CHIU, S. K. 2015. AP4 activates cell migration and EMT mediated by p53 in MDA-MB-231 breast carcinoma cells. *Mol Cell Biochem,* 407**,** 57-68.

CHENG, W., LIU, T., WAN, X., GAO, Y. & WANG, H. 2012. MicroRNA-199a targets CD44 to suppress the tumorigenicity and multidrug resistance of ovarian cancer-initiating cells. *Febs j,* 279**,** 2047-59.

CHENG, W. C., CHUNG, I. F., CHEN, C. Y., SUN, H. J., FEN, J. J., TANG, W. C., CHANG, T. Y., WONG, T. T. & WANG, H. W. 2014. DriverDB: an exome sequencing database for cancer driver gene identification. *Nucleic Acids Res,* 42**,** D1048-54.

CHIA, S. Y., MILAS, M., REDDY, S. K., SIPERSTEIN, A., SKUGOR, M., BRAINARD, J. & GUPTA, M. K. 2007. Thyroid-stimulating hormone receptor messenger ribonucleic acid measurement in blood as a marker for circulating thyroid cancer cells and its role in the preoperative diagnosis of thyroid cancer. *J Clin Endocrinol Metab,* 92**,** 468-75.

CHINNAM, M. & GOODRICH, D. W. 2011. RB1, development, and cancer. *Curr Top Dev Biol,* 94**,** 129-69.

CHO, H. D., LEE, J. E., JUNG, H. Y., OH, M. H., LEE, J. H., JANG, S. H., KIM, K. J., HAN, S. W., KIM, S. Y., KIM, H. J., BAE, S. B. & LEE, H. J. 2015. Loss of Tumor Suppressor ARID1A Protein Expression Correlates with Poor Prognosis in Patients with Primary Breast Cancer. *J Breast Cancer,* 18**,** 339-46.

CHO, W. C. 2007. OncomiRs: the discovery and progress of microRNAs in cancers. *Mol Cancer,* 6**,** 60.

CHOI, W. I., JEON, B. N., YOON, J. H., KOH, D. I., KIM, M. H., YU, M. Y., LEE, K. M., KIM, Y., KIM, K., HUR, S. S., LEE, C. E., KIM, K. S. & HUR, M. W. 2013. The proto-oncoprotein FBI-1 interacts with MBD3 to recruit the Mi-2/NuRD-HDAC complex and BCoR and to silence p21WAF/CDKN1A by DNA methylation. *Nucleic Acids Res,* 41**,** 6403-20.

CHOO, A., PALLADINETTI, P., HOLMES, T., BASU, S., SHEN, S., LOCK, R. B., O'BRIEN, T. A., SYMONDS, G. & DOLNIKOV, A. 2008. siRNA targeting the IRF2 transcription factor inhibits leukaemic cell growth. *Int J Oncol,* 33**,** 175-83.

CHOONG, L. Y., LIM, S., CHONG, P. K., WONG, C. Y., SHAH, N. & LIM, Y. P. 2010. Proteome-wide profiling of the MCF10AT breast cancer progression model. *PLoS One,* 5**,** e11030.

CHRISTOFK, H. R., VANDER HEIDEN, M. G., HARRIS, M. H., RAMANATHAN, A., GERSZTEN, R. E., WEI, R., FLEMING, M. D., SCHREIBER, S. L. & CANTLEY, L. C. 2008. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature,* 452**,** 230-3.

CHUNG, J. H. & BUNZ, F. 2013. A loss-of-function mutation in PTCH1 suggests a role for autocrine hedgehog signaling in colorectal tumorigenesis. *Oncotarget,* 4**,** 2208-11.

CLARK, T. G., CONWAY, S. J., SCOTT, I. C., LABOSKY, P. A., WINNIER, G., BUNDY, J., HOGAN, B. L. & GREENSPAN, D. S. 1999. The mammalian Tolloid-like 1 gene, Tll1, is necessary for normal septation and positioning of the heart. *Development,* 126**,** 2631-42.

COARFA, C., FISKUS, W., EEDUNURI, V. K., RAJAPAKSHE, K., FOLEY, C., CHEW, S. A., SHAH, S. S., GENG, C., SHOU, J., MOHAMED, J. S., O'MALLEY, B. W. & MITSIADES, N. 2015. Comprehensive proteomic profiling identifies the androgen receptor axis and other signaling pathways as targets of microRNAs suppressed in metastatic prostate cancer. *Oncogene*.

COKER, H. & BROCKDORFF, N. 2014. SMCHD1 accumulates at DNA damage sites and facilitates the repair of DNA double-strand breaks. *J Cell Sci,* 127**,** 1869-74.

COLLART, C., CHRISTOV, C. P., SMITH, J. C. & KRUDE, T. 2011. The midblastula transition defines the onset of Y RNA-dependent DNA replication in Xenopus laevis. *Mol Cell Biol,* 31**,** 3857-70.

COMAI, G., BOUTET, A., NEIRIJNCK, Y. & SCHEDL, A. 2010. Expression patterns of the Wtx/Amer gene family during mouse embryonic development. *Dev Dyn,* 239**,** 1867-78.

COMAN, D. 1944. Decreased mutal adhesiveness: a property of cells from squamous cell carcinomas. *Cancer Res,* 4**,** 625.

COOLEN, M., KATZ, S. & BALLY-CUIF, L. 2013. miR-9: a versatile regulator of neurogenesis. *Front Cell Neurosci,* 7**,** 220.

COSTELLO, M., PUGH, T. J., FENNELL, T. J., STEWART, C., LICHTENSTEIN, L., MELDRIM, J. C., FOSTEL, J. L., FRIEDRICH, D. C., PERRIN, D., DIONNE, D., KIM, S., GABRIEL, S. B., LANDER, E. S., FISHER, S. & GETZ, G. 2013. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res,* 41**,** e67.

COURTOIS-COX, S., JONES, S. L. & CICHOWSKI, K. 2008. Many roads lead to oncogene-induced senescence. *Oncogene,* 27**,** 2801-9.

CRESCENZO, R. & INGHIRAMI, G. 2015. Anaplastic lymphoma kinase inhibitors. *Curr Opin Pharmacol,* 23**,** 39-44.

CUI, C., ELSAM, T., TIAN, Q., SEYKORA, J. T., GRACHTCHOUK, M., DLUGOSZ, A. & TSENG, H. 2004. Gli proteins up-regulate the expression of basonuclin in Basal cell carcinoma. *Cancer Res,* 64**,** 5651-8.

CUNNINGHAM, F., AMODE, M. R., BARRELL, D., BEAL, K., BILLIS, K., BRENT, S., CARVALHO-SILVA, D., CLAPHAM, P., COATES, G., FITZGERALD, S., GIL, L., GIRÓN, C. G., GORDON, L., HOURLIER, T., HUNT, S. E., JANACEK, S. H., JOHNSON, N., JUETTEMANN, T., KÄHÄRI, A. K., KEENAN, S., MARTIN, F. J., MAUREL, T., MCLAREN, W., MURPHY, D. N., NAG, R., OVERDUIN, B., PARKER, A., PATRICIO, M., PERRY, E., PIGNATELLI, M., RIAT, H. S., SHEPPARD, D., TAYLOR, K., THORMANN, A., VULLO, A., WILDER, S. P., ZADISSA, A., AKEN, B. L., BIRNEY, E., HARROW, J., KINSELLA, R., MUFFATO, M., RUFFIER, M., SEARLE, S. M. J., SPUDICH, G., TREVANION, S. J., YATES, A., ZERBINO, D. R. & FLICEK, P. 2015. Ensembl 2015.

CURAT, C. A. & VOGEL, W. F. 2002. Discoidin domain receptor 1 controls growth and adhesion of mesangial cells. *J Am Soc Nephrol,* 13**,** 2648-56.

DAHLIN, A. M., HOLLEGAARD, M. V., WIBOM, C., ANDERSSON, U., HOUGAARD, D. M., DELTOUR, I., HJALMARS, U. & MELIN, B. 2015. CCND2, CTNNB1, DDX3X, GLI2, SMARCA4, MYC, MYCN, PTCH1, TP53, and MLL2 gene variants and risk of childhood medulloblastoma. *J Neurooncol,* 125**,** 75-8.

DAHMANE, N., LEE, J., ROBINS, P., HELLER, P. & RUIZ I ALTABA, A. 1997. Activation of the transcription factor Gli1 and the Sonic hedgehog signalling pathway in skin tumours. *Nature,* 389**,** 876-81.

DAKHEL, S., PADILLA, L., ADAN, J., MASA, M., MARTINEZ, J. M., ROQUE, L., COLL, T., HERVAS, R., CALVIS, C., MESSEGUER, R., MITJANS, F. & HERNANDEZ, J. L. 2014. S100P antibody-mediated therapy as a new promising strategy for the treatment of pancreatic cancer. *Oncogenesis,* 3**,** e92.

DAL MOLIN, M., HONG, S. M., HEBBAR, S., SHARMA, R., SCRIMIERI, F., DE WILDE, R. F., MAYO, S. C., GOGGINS, M., WOLFGANG, C. L., SCHULICK, R. D., LIN, M. T., ESHLEMAN, J. R., HRUBAN, R. H., MAITRA, A. & MATTHAEI, H. 2012. Loss of expression of the SWI/SNF chromatin remodeling subunit BRG1/SMARCA4 is frequently observed in intraductal papillary mucinous neoplasms of the pancreas. *Hum Pathol,* 43**,** 585-91.

DALM, S., SIEUWERTS, A., LOOK, M., MELIS, M., VAN DEURZEN, C., FOEKENS, J., DE JONG, M. & MARTENS, J. 2015. Clinical relevance of targeting the gastrin releasing peptide receptor, somatostatin receptor 2 or chemokine c-x-c motif 4 in breast cancer for imaging and therapy. *J Nucl Med*.

DANG, C. V., RESAR, L. M., EMISON, E., KIM, S., LI, Q., PRESCOTT, J. E., WONSEY, D. & ZELLER, K. 1999. Function of the c-Myc oncogenic transcription factor. *Exp Cell Res,* 253**,** 63-77.

DAVIDSON, B., TROPE, C. G. & REICH, R. 2014. The Role of the Tumor Stroma in Ovarian Cancer. *Front Oncol,* 4**,** 104.

DAVIES, K. D. & DOEBELE, R. C. 2013. Molecular pathways: ROS1 fusion proteins in cancer. *Clin Cancer Res,* 19**,** 4040-5.

DAXINGER, L., TAPSCOTT, S. J. & VAN DER MAAREL, S. M. 2015. Genetic and epigenetic contributors to FSHD. *Curr Opin Genet Dev,* 33**,** 56-61.

DAY, B. W., STRINGER, B. W., AL-EJEH, F., TING, M. J., WILSON, J., ENSBEY, K. S., JAMIESON, P. R., BRUCE, Z. C., LIM, Y. C., OFFENHAUSER, C., CHARMSAZ, S., COOPER, L. T., ELLACOTT, J. K., HARDING, A., LEVEQUE, L., INGLIS, P., ALLAN, S., WALKER, D. G., LACKMANN, M., OSBORNE, G., KHANNA, K. K., REYNOLDS, B. A., LICKLITER, J. D. & BOYD, A. W. 2013. EphA3 maintains tumorigenicity and is a therapeutic target in glioblastoma multiforme. *Cancer Cell,* 23**,** 238-48.

DE ARRAS, L., LAWS, R., LEACH, S. M., PONTIS, K., FREEDMAN, J. H., SCHWARTZ, D. A. & ALPER, S. 2014. Comparative genomics RNAi screen identifies Eftud2 as a novel regulator of innate immunity. *Genetics,* 197**,** 485-96.

DE HELLER-MILEV, M., HUBER, M., PANIZZON, R. & HOHL, D. 2000. Expression of small proline rich proteins in neoplastic and inflammatory skin diseases. *Br J Dermatol,* 143**,** 733-40.

DE KONING, P. J., KUMMER, J. A., DE POOT, S. A., QUADIR, R., BROEKHUIZEN, R., MCGETTRICK, A. F., HIGGINS, W. J., DEVREESE, B., WORRALL, D. M. & BOVENSCHEN, N. 2011. Intracellular serine protease inhibitor SERPINB4 inhibits granzyme M-induced cell death. *PLoS One,* 6**,** e22645.

DEAR, N., MATENA, K., VINGRON, M. & BOEHM, T. 1997. A new subfamily of vertebrate calpains lacking a calmodulin-like domain: implications for calpain regulation and evolution. *Genomics,* 45**,** 175-84.

DELGADO, A. P., BRANDAO, P., CHAPADO, M. J., HAMID, S. & NARAYANAN, R. 2014a. Open reading frames associated with cancer in the dark matter of the human genome. *Cancer Genomics Proteomics,* 11**,** 201-13.

DELGADO, A. P., HAMID, S., BRANDAO, P. & NARAYANAN, R. 2014b. A novel transmembrane glycoprotein cancer biomarker present in the X chromosome. *Cancer Genomics Proteomics,* 11**,** 81-92.

DEMETRICK, D. J., MATSUMOTO, S., HANNON, G. J., OKAMOTO, K., XIONG, Y., ZHANG, H. & BEACH, D. H. 1995. Chromosomal mapping of the genes for the human cell cycle proteins cyclin C (CCNC), cyclin E (CCNE), p21 (CDKN1) and KAP (CDKN3). *Cytogenet Cell Genet,* 69**,** 190-2.

DEMIRAG, G. G., SULLU, Y., GURGENYATAGI, D., OKUMUS, N. O. & YUCEL, I. 2011. Expression of plakophilins (PKP1, PKP2, and PKP3) in gastric cancers. *Diagn Pathol,* 6**,** 1.

DEMIRAG, G. G., SULLU, Y. & YUCEL, I. 2012. Expression of Plakophilins (PKP1, PKP2, and PKP3) in breast cancers. *Med Oncol,* 29**,** 1518-22.

DEMIRCI, H., WORDEN, F., NELSON, C. C., ELNER, V. M. & KAHANA, A. 2015. Efficacy of Vismodegib (Erivedge) for Basal Cell Carcinoma Involving the Orbit and Periocular Area. *Ophthal Plast Reconstr Surg*.

DENG, B., ZHANG, Y., ZHANG, S., WEN, F., MIAO, Y. & GUO, K. 2015. MicroRNA-142-3p inhibits cell proliferation and invasion of cervical cancer cells by targeting FZD7. *Tumour Biol*.

DENSLOW, S. A. & WADE, P. A. 2007. The human Mi-2/NuRD complex and gene regulation. *Oncogene,* 26**,** 5433-8.

DEPEW, M. J., LUFKIN, T. & RUBENSTEIN, J. L. 2002. Specification of jaw subdivisions by Dlx genes. *Science,* 298**,** 381-5.

DEPREZ, M. & UFFER, S. 2009. Clinicopathological features of eyelid skin tumors. A retrospective study of 5504 cases and review of literature. *Am J Dermatopathol,* 31**,** 256-62.

DERKINDEREN, D. J., KOTEN, J. W., WOLTERBEEK, R., BEEMER, F. A., TAN, K. E. & DEN OTTER, W. 1987. Non-ocular cancer in hereditary retinoblastoma survivors and relatives. *Ophthalmic Paediatr Genet,* 8**,** 23-5.

DEZHONG, L., XIAOYI, Z., XIANLIAN, L., HONGYAN, Z., GUOHUA, Z., BO, S., SHENGLEI, Z. & LIAN, Z. 2015. miR-150 is a factor of survival in prostate cancer patients. *J buon,* 20**,** 173-9.

DI GAETANO, N., CITTERA, E., NOTA, R., VECCHI, A., GRIECO, V., SCANZIANI, E., BOTTO, M., INTRONA, M. & GOLAY, J. 2003. Complement activation determines the therapeutic activity of rituximab in vivo. *J Immunol,* 171**,** 1581-7.

DIECI, M. V., SMUTNA, V., SCOTT, V., YIN, G., XU, R., VIELH, P., MATHIEU, M. C., VICIER, C., LAPORTE, M., DRUSCH, F., GUARNERI, V., CONTE, P., DELALOGE, S., LACROIX, L., FROMIGUE, O., ANDRE, F. & LEFEBVRE, C. 2016. Whole exome sequencing of rare aggressive breast cancer histologies. *Breast Cancer Res Treat,* 156**,** 21-32.

DING, L., GETZ, G., WHEELER, D. A., MARDIS, E. R., MCLELLAN, M. D., CIBULSKIS, K., SOUGNEZ, C., GREULICH, H., MUZNY, D. M., MORGAN, M. B., FULTON, L., FULTON, R. S., ZHANG, Q., WENDL, M. C., LAWRENCE, M. S., LARSON, D. E., CHEN, K., DOOLING, D. J., SABO, A., HAWES, A. C., SHEN, H., JHANGIANI, S. N., LEWIS, L. R., HALL, O., ZHU, Y., MATHEW, T., REN, Y., YAO, J., SCHERER, S. E., CLERC, K., METCALF, G. A., NG, B., MILOSAVLJEVIC, A., GONZALEZ-GARAY, M. L., OSBORNE, J. R., MEYER, R., SHI, X., TANG, Y., KOBOLDT, D. C., LIN, L., ABBOTT, R., MINER, T. L., POHL, C., FEWELL, G., HAIPEK, C., SCHMIDT, H., DUNFORD-SHORE, B. H., KRAJA, A., CROSBY, S. D., SAWYER, C. S., VICKERY, T., SANDER, S., ROBINSON, J., WINCKLER, W., BALDWIN, J., CHIRIEAC, L. R., DUTT, A., FENNELL, T., HANNA, M., JOHNSON, B. E., ONOFRIO, R. C., THOMAS, R. K., TONON, G., WEIR, B. A., ZHAO, X., ZIAUGRA, L., ZODY, M. C., GIORDANO, T., ORRINGER, M. B., ROTH, J. A., SPITZ, M. R., WISTUBA, II, OZENBERGER, B., GOOD, P. J., CHANG, A. C., BEER, D. G., WATSON, M. A., LADANYI, M., BRODERICK, S., YOSHIZAWA, A., TRAVIS, W. D., PAO, W., PROVINCE, M. A., WEINSTOCK, G. M., VARMUS, H. E., GABRIEL, S. B., LANDER, E. S., GIBBS, R. A., MEYERSON, M. & WILSON, R. K. 2008. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature,* 455**,** 1069-75.

DOHERTY, J. A., ROSSING, M. A., CUSHING-HAUGEN, K. L., CHEN, C., VAN DEN BERG, D. J., WU, A. H., PIKE, M. C., NESS, R. B., MOYSICH, K., CHENEVIX-TRENCH, G., BEESLEY, J., WEBB, P. M., CHANG-CLAUDE, J., WANG-GOHRKE, S., GOODMAN, M. T., LURIE, G., THOMPSON, P. J.,

CARNEY, M. E., HOGDALL, E., KJAER, S. K., HOGDALL, C., GOODE, E. L., CUNNINGHAM, J. M., FRIDLEY, B. L., VIERKANT, R. A., BERCHUCK, A., MOORMAN, P. G., SCHILDKRAUT, J. M., PALMIERI, R. T., CRAMER, D. W., TERRY, K. L., YANG, H. P., GARCIA-CLOSAS, M., CHANOCK, S., LISSOWSKA, J., SONG, H., PHAROAH, P. D., SHAH, M., PERKINS, B., MCGUIRE, V., WHITTEMORE, A. S., DI CIOCCIO, R. A., GENTRY-MAHARAJ, A., MENON, U., GAYTHER, S. A., RAMUS, S. J., ZIOGAS, A., BREWSTER, W., ANTON-CULVER, H. & PEARCE, C. L. 2010. ESR1/SYNE1 polymorphism and invasive epithelial ovarian cancer risk: an Ovarian Cancer Association Consortium study. *Cancer Epidemiol Biomarkers Prev,* 19**,** 245-50.

DONG, C., SLATTERY, M. J., LIANG, S. & PENG, H. H. 2005. Melanoma cell extravasation under flow conditions is modulated by leukocytes and endogenously produced interleukin 8. *Mol Cell Biomech,* 2**,** 145-59.

DONG, S., KANNO, T., YAMAKI, A., KOJIMA, T., SHIRAIWA, M., KAWADA, A., MECHIN, M. C., CHAVANAS, S., SERRE, G., SIMON, M. & TAKAHARA, H. 2006. NF-Y and Sp1/Sp3 are involved in the transcriptional regulation of the peptidylarginine deiminase type III gene (PADI3) in human keratinocytes. *Biochem J,* 397**,** 449-59.

DONLEY, N. & THAYER, M. J. 2013. DNA replication timing, genome stability and cancer: late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin Cancer Biol,* 23**,** 80-9.

DORAJOO, R., SUN, Y., HAN, Y., KE, T., BURGER, A., CHANG, X., LOW, H. Q., GUAN, W., LEMAITRE, R. N., KHOR, C. C., YUAN, J. M., KOH, W. P., ONG, C. N., TAI, E. S., LIU, J., VAN DAM, R. M., HENG, C. K. & FRIEDLANDER, Y. 2015. A genome-wide association study of n-3 and n-6 plasma fatty acids in a Singaporean Chinese population. *Genes Nutr,* 10**,** 53.

DOXANAS, M. T. & GREEN, W. R. 1984. Sebaceous gland carcinoma. Review of 40 cases. *Arch Ophthalmol,* 102**,** 245-9.

DRAKE, J. W., CHARLESWORTH, B., CHARLESWORTH, D. & CROW, J. F. 1998. Rates of spontaneous mutation. *Genetics,* 148**,** 1667-86.

DU, F., FENG, Y., FANG, J. & YANG, M. 2015. MicroRNA-143 enhances chemosensitivity of Quercetin through autophagy inhibition via target GABARAPL1 in gastric cancer cells. *Biomed Pharmacother,* 74**,** 169-77.

DU, J. L., WEI, L. X. & WANG, Y. L. 2011. [Expression and clinicopathologic significance of GPC3 and other antibodies in well-differentiated hepatocellular carcinoma]. *Zhonghua Bing Li Xue Za Zhi,* 40**,** 11-6.

DU, W. W., YANG, W. & YEE, A. J. 2013. Roles of versican in cancer biology--tumorigenesis, progression and metastasis. *Histol Histopathol,* 28**,** 701-13.

DUGGINENI, S., MITRA, S., NOBERINI, R., HAN, X., LIN, N., XU, Y., TIAN, W., AN, J., PASQUALE, E. B. & HUANG, Z. 2013. Design, synthesis and characterization of novel small molecular inhibitors of ephrin-B2 binding to EphB4. *Biochem Pharmacol,* 85**,** 507-13.

EASTON, D. F., PHAROAH, P. D., ANTONIOU, A. C., TISCHKOWITZ, M., TAVTIGIAN, S. V., NATHANSON, K. L., DEVILEE, P., MEINDL, A., COUCH, F. J., SOUTHEY, M., GOLDGAR, D. E., EVANS, D. G., CHENEVIX-TRENCH, G., RAHMAN, N., ROBSON, M., DOMCHEK, S. M. & FOULKES, W. D. 2015. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med,* 372**,** 2243-57.

EDGE SB, B. D., COMPTON CC, FRITZ AG, GREENE FL, TROTTI A 2010. *Cancer Staging Manual (7th edition), pp 299–344,* New York, Springer.

EELES, R. A., OLAMA, A. A., BENLLOCH, S., SAUNDERS, E. J., LEONGAMORNLERT, D. A., TYMRAKIEWICZ, M., GHOUSSAINI, M., LUCCARINI, C., DENNIS, J., JUGURNAUTH-LITTLE, S., DADAEV, T., NEAL, D. E., HAMDY, F. C., DONOVAN, J. L., MUIR, K., GILES, G. G., SEVERI, G., WIKLUND, F., GRONBERG, H., HAIMAN, C. A., SCHUMACHER, F., HENDERSON, B. E., LE MARCHAND, L., LINDSTROM, S., KRAFT, P., HUNTER, D. J., GAPSTUR, S., CHANOCK, S. J., BERNDT, S. I., ALBANES, D., ANDRIOLE, G., SCHLEUTKER, J., WEISCHER, M., CANZIAN, F., RIBOLI, E., KEY, T. J., TRAVIS, R. C., CAMPA, D., INGLES, S. A., JOHN, E. M., HAYES, R. B., PHAROAH, P. D., PASHAYAN, N., KHAW, K. T., STANFORD, J. L., OSTRANDER, E. A., SIGNORELLO, L. B., THIBODEAU, S. N., SCHAID, D., MAIER, C., VOGEL, W., KIBEL, A. S., CYBULSKI, C., LUBINSKI, J., CANNON-ALBRIGHT, L., BRENNER, H., PARK, J. Y., KANEVA, R., BATRA, J., SPURDLE, A. B., CLEMENTS, J. A., TEIXEIRA, M. R., DICKS, E., LEE, A., DUNNING, A. M., BAYNES, C., CONROY, D., MARANIAN, M. J., AHMED, S., GOVINDASAMI, K., GUY, M., WILKINSON, R. A., SAWYER, E. J., MORGAN, A., DEARNALEY, D. P., HORWICH, A., HUDDART, R. A., KHOO, V. S., PARKER, C. C., VAN AS, N. J., WOODHOUSE, C. J., THOMPSON, A., DUDDERIDGE, T., OGDEN, C., COOPER, C. S., LOPHATANANON, A., COX, A., SOUTHEY, M. C., HOPPER, J. L., ENGLISH, D. R., ALY, M., ADOLFSSON, J., XU, J., ZHENG, S. L., YEAGER, M., KAAKS, R., DIVER, W. R., GAUDET, M. M., STERN, M. C., CORRAL, R., et al. 2013. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat Genet,* 45**,** 385-91, 391e1-2.

ELKAMHAWY, A., PARK, J. E., CHO, N. C., SIM, T., PAE, A. N. & ROH, E. J. 2015. Discovery of a broad spectrum antiproliferative agent with selectivity for DDR1 kinase: cell line-based assay, kinase panel, molecular docking, and toxicity studies. *J Enzyme Inhib Med Chem***,** 1-9.

EMAN, R. M., HOORNTJE, E. T., ONER, F. C., KRUYT, M. C., DHERT, W. J. & ALBLAS, J. 2014. CXCL12/stromal-cell-derived factor-1 effectively replaces endothelial progenitor cells to induce vascularized ectopic bone. *Stem Cells Dev,* 23**,** 2950-8.

ESNAULT, C., MAESTRE, J. & HEIDMANN, T. 2000. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet,* 24**,** 363-7.

EWING, J. P. O. P. A. C. U. M. C. 1919. *Neoplastic Diseases: a text-book on tumors, etc,* Philadelphia ; London, W.B. Saunders Co.

FADLOUN, A., KOBI, D., POINTUD, J. C., INDRA, A. K., TELETIN, M., BOLE-FEYSOT, C., TESTONI, B., MANTOVANI, R., METZGER, D., MENGUS, G. & DAVIDSON, I. 2007. The TFIID subunit TAF4 regulates keratinocyte proliferation and has cell-autonomous and non-cell-autonomous tumour suppressor activity in mouse epidermis. *Development,* 134**,** 2947-58.

FARNDON, P. A., DEL MASTRO, R. G., EVANS, D. G. & KILPATRICK, M. W. 1992. Location of gene for Gorlin syndrome. *Lancet,* 339**,** 581-2.

FENG, J., YANG, Y., ZHANG, P., WANG, F., MA, Y., QIN, H. & WANG, Y. 2014. miR-150 functions as a tumour suppressor in human colorectal cancer by targeting c-Myb. *J Cell Mol Med,* 18**,** 2125-34.

FENG, R., CHEN, X., YU, Y., SU, L., YU, B., LI, J., CAI, Q., YAN, M., LIU, B. & ZHU, Z. 2010. miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer Lett,* 298**,** 50-63.

FENG, Y. & WALSH, C. A. 2004. The many faces of filamin: a versatile molecular scaffold for cell motility and signalling. *Nat Cell Biol,* 6**,** 1034-8.

FERGUSON, B. D., TRETIAKOVA, M. S., LINGEN, M. W., GILL, P. S. & SALGIA, R. 2014. Expression of the EPHB4 receptor tyrosine kinase in head and neck and renal malignancies--implications for solid tumors and potential for therapeutic inhibition. *Growth Factors,* 32**,** 202-6.

FILMUS, J., CAPURRO, M. & RAST, J. 2008. Glypicans. *Genome Biol,* 9**,** 224.

FIORENTINO, F. P., TOKGUN, E., SOLE-SANCHEZ, S., GIAMPAOLO, S., TOKGUN, O., JAUSET, T., KOHNO, T., PERUCHO, M., SOUCEK, L. & YOKOTA, J. 2016. Growth suppression by MYC inhibition in small cell lung cancer cells with TP53 and RB1 inactivation. *Oncotarget*.

FIRNHABER, J. M. 2012. Diagnosis and treatment of Basal cell and squamous cell carcinoma. *Am Fam Physician,* 86**,** 161-8.

FISCHER-KESO, R., BREUNINGER, S., HOFMANN, S., HENN, M., ROHRIG, T., STROBEL, P., STOECKLIN, G. & HOFMANN, I. 2014. Plakophilins 1 and 3 bind to FXR1 and thereby influence the mRNA stability of desmosomal proteins. *Mol Cell Biol,* 34**,** 4244-56.

FLOHIL, S. C., SEUBRING, I., VAN ROSSUM, M. M., COEBERGH, J. W., DE VRIES, E. & NIJSTEN, T. 2013. Trends in Basal cell carcinoma incidence rates: a 37-year Dutch observational study. *J Invest Dermatol,* 133**,** 913-8.

FLORA, A., KLISCH, T. J., SCHUSTER, G. & ZOGHBI, H. Y. 2009. Deletion of Atoh1 disrupts Sonic Hedgehog signaling in the developing cerebellum and prevents medulloblastoma. *Science.* United States.

FORBES, S. A., BINDAL, N., BAMFORD, S., COLE, C., KOK, C. Y., BEARE, D., JIA, M., SHEPHERD, R., LEUNG, K., MENZIES, A., TEAGUE, J. W., CAMPBELL, P. J., STRATTON, M. R. & FUTREAL, P. A. 2011. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* England.

FORMENTINI, A., HENNE-BRUNS, D. & KORNMANN, M. 2004. Thymidylate synthase expression and prognosis of patients with gastrointestinal cancers receiving adjuvant chemotherapy: a review. *Langenbecks Arch Surg,* 389**,** 405-13.

FOSSET, C., CHAUVEAU, M. J., GUILLON, B., CANAL, F., DRAPIER, J. C. & BOUTON, C. 2006. RNA silencing of mitochondrial m-Nfs1 reduces Fe-S enzyme activity both in mitochondria and cytosol of mammalian cells. *J Biol Chem,* 281**,** 25398-406.

FRANCIS, H., DEMORROW, S., VENTER, J., ONORI, P., WHITE, M., GAUDIO, E., FRANCIS, T., GREENE, J. F., JR., TRAN, S., MEININGER, C. J. & ALPINI, G. 2012. Inhibition of histidine decarboxylase ablates the autocrine tumorigenic effects of histamine in human cholangiocarcinoma. *Gut,* 61**,** 753-64.

FRANCO-CERECEDA, A., GENNARI, C., NAMI, R., AGNUSDEI, D., PERNOW, J., LUNDBERG, J. M. & FISCHER, J. A. 1987. Cardiovascular effects of calcitonin gene-related peptides I and II in man. *Circ Res,* 60**,** 393-7.

FRANKLIN, R. B., MA, J., ZOU, J., GUAN, Z., KUKOYI, B. I., FENG, P. & COSTELLO, L. C. 2003. Human ZIP1 is a major zinc uptake transporter for the accumulation of zinc in prostate cells. *J Inorg Biochem,* 96**,** 435-42.

FREYTAG, S. O., STRICKER, H., MOVSAS, B. & KIM, J. H. 2007. Prostate Cancer Gene Therapy Clinical Trials. *Molecular Therapy,* 15**,** 1042-1052.

FRISTEDT, R., ELEBRO, J., GABER, A., JONSSON, L., HEBY, M., YUDINA, Y., NODIN, B., UHLEN, M., EBERHARD, J. & JIRSTROM, K. 2014a. Reduced expression of the polymeric immunoglobulin receptor in pancreatic and periampullary adenocarcinoma signifies tumour progression and poor prognosis. *PLoS One,* 9**,** e112728.

FRISTEDT, R., GABER, A., HEDNER, C., NODIN, B., UHLEN, M., EBERHARD, J. & JIRSTROM, K. 2014b. Expression and prognostic significance of the polymeric immunoglobulin receptor in esophageal and gastric adenocarcinoma. *J Transl Med,* 12**,** 83.

FU, J., QIN, L., HE, T., QIN, J., HONG, J., WONG, J., LIAO, L. & XU, J. 2011. The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis. *Cell Res,* 21**,** 275-89.

FU, X. C. & XIANG, W. P. 2012. [Testicular CR16 and spermatogenesis]. *Zhonghua Nan Ke Xue,* 18**,** 1032-5.

FUJIWARA, T., HARIGAE, H., OKITSU, Y., TAKAHASHI, S., YOKOYAMA, H., YAMADA, M. F., ISHIZAWA, K., KAMEOKA, J., KAKU, M. & SASAKI, T. 2006. Expression analyses and transcriptional regulation of mouse nucleolar spindle-associated protein gene in erythroid cells: essential role of NF-Y. *Br J Haematol,* 135**,** 583-90.

FUKAMI, S., RIEMENSCHNEIDER, M. J., KOHNO, M. & STEIGER, H. J. 2016. Expression and gene doses changes of the p53-regulator PPM1D in meningiomas: a role in meningioma progression? *Brain Tumor Pathol*.

GAILANI, M. R., STÅHLE-BÄCKDAHL, M., LEFFELL, D. J., GLYN, M., ZAPHIROPOULOS, P. G., UNDÉN, A. B., DEAN, M., BRASH, D. E., BALE, A. E. & TOFTGÅRD, R. 1996. The role of the human homologue of Drosophila patched in sporadic basal cell carcinomas. *Nature Genetics,* 14**,** 78-81.

GAO, S. L., WANG, L. Z., LIU, H. Y., LIU, D. L., XIE, L. M. & ZHANG, Z. W. 2014. miR-200a inhibits tumor proliferation by targeting AP-2gamma in neuroblastoma cells. *Asian Pac J Cancer Prev,* 15**,** 4671-6.

GAO, Y., FENG, Y., SHEN, J. K., LIN, M., CHOY, E., COTE, G. M., HARMON, D. C., MANKIN, H. J., HORNICEK, F. J. & DUAN, Z. 2015. CD44 is a direct target of miR-199a-3p and contributes to aggressive progression in osteosarcoma. *Sci Rep,* 5**,** 11365.

GARCIA, E., MACHESKY, L. M., JONES, G. E. & ANTON, I. M. 2014. WIP is necessary for matrix invasion by breast cancer cells. *Eur J Cell Biol,* 93**,** 413-23.

GARCIA-CLOSAS, M., COUCH, F. J., LINDSTROM, S., MICHAILIDOU, K., SCHMIDT, M. K., BROOK, M. N., ORR, N., RHIE, S. K., RIBOLI, E., FEIGELSON, H. S., LE MARCHAND, L., BURING, J. E., ECCLES, D., MIRON, P., FASCHING, P. A., BRAUCH, H., CHANG-CLAUDE, J., CARPENTER, J., GODWIN, A. K., NEVANLINNA, H., GILES, G. G., COX, A., HOPPER, J. L., BOLLA, M. K., WANG, Q., DENNIS, J., DICKS, E., HOWAT, W. J., SCHOOF, N., BOJESEN, S. E., LAMBRECHTS, D., BROEKS, A., ANDRULIS, I. L., GUENEL, P., BURWINKEL, B., SAWYER, E. J., HOLLESTELLE, A., FLETCHER, O., WINQVIST, R., BRENNER, H., MANNERMAA, A., HAMANN, U., MEINDL, A., LINDBLOM, A., ZHENG, W., DEVILLEE, P., GOLDBERG, M. S., LUBINSKI, J., KRISTENSEN, V., SWERDLOW, A., ANTON-CULVER, H., DORK, T., MUIR, K., MATSUO,

K., WU, A. H., RADICE, P., TEO, S. H., SHU, X. O., BLOT, W., KANG, D., HARTMAN, M., SANGRAJRANG, S., SHEN, C. Y., SOUTHEY, M. C., PARK, D. J., HAMMET, F., STONE, J., VEER, L. J., RUTGERS, E. J., LOPHATANANON, A., STEWART-BROWN, S., SIRIWANARANGSAN, P., PETO, J., SCHRAUDER, M. G., EKICI, A. B., BECKMANN, M. W., DOS SANTOS SILVA, I., JOHNSON, N., WARREN, H., TOMLINSON, I., KERIN, M. J., MILLER, N., MARME, F., SCHNEEWEISS, A., SOHN, C., TRUONG, T., LAURENT-PUIG, P., KERBRAT, P., NORDESTGAARD, B. G., NIELSEN, S. F., FLYGER, H., MILNE, R. L., PEREZ, J. I., MENENDEZ, P., MULLER, H., ARNDT, V., STEGMAIER, C., LICHTNER, P., LOCHMANN, M., JUSTENHOVEN, C., et al. 2013. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet,* 45**,** 392-8, 398e1-2.

GEISSE, J., CARO, I., LINDHOLM, J., GOLITZ, L., STAMPONE, P. & OWENS, M. 2004. Imiquimod 5% cream for the treatment of superficial basal cell carcinoma: results from two phase III, randomized, vehicle-controlled studies. *J Am Acad Dermatol,* 50**,** 722-33.

GILL, H. S., MOSCATO, E. E., CHANG, A. L., SOON, S. & SILKISS, R. Z. 2013. Vismodegib for periocular and orbital basal cell carcinoma. *JAMA Ophthalmol,* 131**,** 1591-4.

GLUNDE, K., PENET, M. F., JIANG, L., JACOBS, M. A. & BHUJWALLA, Z. M. 2015. Choline metabolism-based molecular diagnosis of cancer: an update. *Expert Rev Mol Diagn,* 15**,** 735-47.

GNANAPRAGASAM, M. N., SCARSDALE, J. N., AMAYA, M. L., WEBB, H. D., DESAI, M. A., WALAVALKAR, N. M., WANG, S. Z., ZU ZHU, S., GINDER, G. D. & WILLIAMS, D. C., JR. 2011. p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc Natl Acad Sci U S A,* 108**,** 7487-92.

GOLAN, T., YANIV, A., BAFICO, A., LIU, G. & GAZIT, A. 2004. The human Frizzled 6 (HFz6) acts as a negative regulator of the canonical Wnt. beta-catenin signaling cascade. *J Biol Chem,* 279**,** 14879-88.

GONG, J. & HUO, J. 2015. New insights into the mechanism of F-box proteins in colorectal cancer (Review). *Oncol Rep,* 33**,** 2113-20.

GONNISSEN, A., ISEBAERT, S. & HAUSTERMANS, K. 2015. Targeting the Hedgehog signaling pathway in cancer: beyond Smoothened. *Oncotarget,* 6**,** 13899-913.

GONZALEZ-PEREZ, A., DEU-PONS, J. & LOPEZ-BIGAS, N. 2012. Improving the prediction of the functional impact of cancer mutations by baseline tolerance transformation. *Genome Med,* 4**,** 89.

GONZALEZ-PEREZ, A., PEREZ-LLAMAS, C., DEU-PONS, J., TAMBORERO, D., SCHROEDER, M. P., JENE-SANZ, A., SANTOS, A. & LOPEZ-BIGAS, N. 2013. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods,* 10**,** 1081-2.

GORBACHEV, A. V. & FAIRCHILD, R. L. 2014. Regulation of chemokine expression in the tumor microenvironment. *Crit Rev Immunol,* 34**,** 103-20.

GORDON, C. A., GULZAR, Z. G. & BROOKS, J. D. 2015. NUSAP1 expression is upregulated by loss of RB1 in prostate cancer cells. *Prostate,* 75**,** 517-26.

GREGORY, P. A., BERT, A. G., PATERSON, E. L., BARRY, S. C., TSYKIN, A., FARSHID, G., VADAS, M. A., KHEW-GOODALL, Y. & GOODALL, G. J.

2008. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol,* 10**,** 593-601.

GRIMSHAW, M. J., NAYLOR, S. & BALKWILL, F. R. 2002. Endothelin-2 is a hypoxia-induced autocrine survival factor for breast tumor cells. *Mol Cancer Ther,* 1**,** 1273-81.

GRISHINA, I. B. 2015. Mini-review: Does Notch promote or suppress cancer? New findings and old controversies. *Am J Clin Exp Urol,* 3**,** 24-7.

GROLIMUND, L., AEBY, E., HAMELIN, R., ARMAND, F., CHIAPPE, D., MONIATTE, M. & LINGNER, J. 2013. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun,* 4**,** 2848.

GROS-LOUIS, F., DUPRE, N., DION, P., FOX, M. A., LAURENT, S., VERREAULT, S., SANES, J. R., BOUCHARD, J. P. & ROULEAU, G. A. 2007. Mutations in SYNE1 lead to a newly discovered form of autosomal recessive cerebellar ataxia. *Nat Genet,* 39**,** 80-5.

GRUBER, J., SEE TOO, W. C., WONG, M. T., LAVIE, A., MCSORLEY, T. & KONRAD, M. 2012. Balance of human choline kinase isoforms is critical for cell cycle regulation: implications for the development of choline kinase-targeted cancer therapy. *Febs j,* 279**,** 1915-28.

GRUNSTEIN, M. 1990. Histone function in transcription. *Annu Rev Cell Biol,* 6**,** 643-78.

GUDMANN, N. S., HANSEN, N. U., JENSEN, A. C., KARSDAL, M. A. & SIEBUHR, A. S. 2015. Biological relevance of citrullinations: diagnostic, prognostic and therapeutic options. *Autoimmunity,* 48**,** 73-9.

GULZAR, Z. G., MCKENNEY, J. K. & BROOKS, J. D. 2013. Increased expression of NuSAP in recurrent prostate cancer is mediated by E2F1. *Oncogene,* 32**,** 70-7.

GUNDA, V., COGDILL, A. P., BERNASCONI, M. J., WARGO, J. A. & PARANGI, S. 2013. Potential role of 5-aza-2'-deoxycytidine induced MAGE-A4 expression in immunotherapy for anaplastic thyroid cancer. *Surgery,* 154**,** 1456-62; discussion 1462.

GUO, H. R., YU, H. S., HU, H. & MONSON, R. R. 2001. Arsenic in drinking water and skin cancers: cell-type specificity (Taiwan, ROC). *Cancer Causes Control,* 12**,** 909-16.

GUO, M., ZHANG, X., WANG, G., SUN, J., JIANG, Z., KHADARIAN, K., YU, S., ZHAO, Y., XIE, C., ZHANG, K., ZHU, M., SHEN, H., LIN, Z., JIANG, C., SHEN, J. & ZHENG, Y. 2015. miR-603 promotes glioma cell growth via Wnt/beta-catenin pathway by inhibiting WIF1 and CTNNBIP1. *Cancer Lett,* 360**,** 76-86.

GUO, Y., ZHANG, W., GIROUX, C., CAI, Y., EKAMBARAM, P., DILLY, A. K., HSU, A., ZHOU, S., MADDIPATI, K. R., LIU, J., JOSHI, S., TUCKER, S. C., LEE, M. J. & HONN, K. V. 2011. Identification of the orphan G protein-coupled receptor GPR31 as a receptor for 12-(S)-hydroxyeicosatetraenoic acid. *J Biol Chem,* 286**,** 33832-40.

HAAK-FRENDSCHO, M., DARVAS, Z., HEGYESI, H., KARPATI, S., HOFFMAN, R. L., LASZLO, V., BENCSATH, M., SZALAI, C., FURESZ, J., TIMAR, J., BATA-CSORGO, Z., SZABAD, G., PIVARCSI, A., PALLINGER, E., KEMENY, L., HORVATH, A., DOBOZY, A. & FALUS, A. 2000. Histidine decarboxylase expression in human melanoma. *J Invest Dermatol,* 115**,** 345-52.

HADLER-OLSEN, E., WINBERG, J. O. & UHLIN-HANSEN, L. 2013. Matrix metalloproteinases in cancer: their value as diagnostic and prognostic markers and therapeutic targets. *Tumour Biol,* 34**,** 2041-51.

HAIDER, A. S., PETERS, S. B., KAPORIS, H., CARDINALE, I., FEI, J., OTT, J., BLUMENBERG, M., BOWCOCK, A. M., KRUEGER, J. G. & CARUCCI, J. A. 2006. Genomic analysis defines a cancer-specific gene expression signature for human squamous cell carcinoma and distinguishes malignant hyperproliferation from benign hyperplasia. *J Invest Dermatol,* 126**,** 869-81.

HAN, Y., KUANG, Y., XUE, X., GUO, X., LI, P., WANG, X., YUAN, B., ZHI, Q. & ZHAO, H. 2014. NLK, a novel target of miR-199a-3p, functions as a tumor suppressor in colorectal cancer. *Biomed Pharmacother,* 68**,** 497-505.

HAN, Z., ZHANG, Y., YANG, Q., LIU, B., WU, J., YANG, C. & JIANG, Y. 2015. miR-497 and miR-34a retard lung cancer growth by co-inhibiting cyclin E1 (CCNE1). *Oncotarget,* 6**,** 13149-63.

HARCHARIK, S., BERNARDO, S., MOSKALENKO, M., PAN, M., SIVENDRAN, M., BELL, H., HALL, L. D., CASTILLO-MARTIN, M., FOX, K., CORDON-CARDO, C., CHANG, R., SIVENDRAN, S., PHELPS, R. G. & SAENGER, Y. 2014. Defining the role of CD2 in disease progression and overall survival among patients with completely resected stage-II to -III cutaneous melanoma. *J Am Acad Dermatol,* 70**,** 1036-44.

HARRIS, H., MILLER, O. J., KLEIN, G., WORST, P. & TACHIBANA, T. 1969. Suppression of malignancy by cell fusion. *Nature,* 223**,** 363-8.

HARRIS, R. S. 2013. Cancer mutation signatures, DNA damage mechanisms, and potential clinical implications. *Genome Med.*

HARTANTO, F. K., KAREN-NG, L. P., VINCENT-CHONG, V. K., ISMAIL, S. M., MUSTAFA, W. M., ABRAHAM, M. T., TAY, K. K. & ZAIN, R. B. 2015. KRT13, FAIM2 and CYP2W1 mRNA expression in oral squamous cell carcinoma patients with risk habits. *Asian Pac J Cancer Prev,* 16**,** 953-8.

HARVEY, J. T. & ANDERSON, R. L. 1982. The management of meibomian gland carcinoma. *Ophthalmic Surg,* 13**,** 56-61.

HASUMI, H., BABA, M., HASUMI, Y., LANG, M., HUANG, Y., OH, H. F., MATSUO, M., MERINO, M. J., YAO, M., ITO, Y., FURUYA, M., IRIBE, Y., KODAMA, T., SOUTHON, E., TESSAROLLO, L., NAGASHIMA, K., HAINES, D. C., LINEHAN, W. M. & SCHMIDT, L. S. 2015. Folliculin-interacting proteins Fnip1 and Fnip2 play critical roles in kidney tumor suppression in cooperation with Flcn. *Proc Natl Acad Sci U S A,* 112**,** E1624-31.

HATA, M., KOIKE, I., OMURA, M., MAEGAWA, J., OGINO, I. & INOUE, T. 2012. Noninvasive and curative radiation therapy for sebaceous carcinoma of the eyelid. *Int J Radiat Oncol Biol Phys,* 82**,** 605-11.

HAYASHI, M., NOMOTO, S., HISHIDA, M., INOKAWA, Y., KANDA, M., OKAMURA, Y., NISHIKAWA, Y., TANAKA, C., KOBAYASHI, D., YAMADA, S., NAKAYAMA, G., FUJII, T., SUGIMOTO, H., KOIKE, M., FUJIWARA, M., TAKEDA, S. & KODERA, Y. 2014. Identification of the collagen type 1 alpha 1 gene (COL1A1) as a candidate survival-related factor associated with hepatocellular carcinoma. *BMC Cancer,* 14**,** 108.

HAYASHI, Y., OSANAI, M. & LEE, G. H. 2015. NOTCH2 signaling confers immature morphology and aggressiveness in human hepatocellular carcinoma cells. *Oncol Rep,* 34**,** 1650-8.

HAYDEN, C. A. & BOSCO, G. 2008. Comparative genomic analysis of novel conserved peptide upstream open reading frames in Drosophila melanogaster and other dipteran species. *BMC Genomics,* 9**,** 61.

HE, J., WANG, F., ZHU, J. H., CHEN, W., CUI, Z. & JIA, W. H. 2015. No association between MTR rs1805087 A > G polymorphism and non-Hodgkin lymphoma susceptibility: evidence from 11 486 subjects. *Leuk Lymphoma,* 56**,** 763-7.

HE, J. M., PU, Y. D., WU, Y. J., QIN, R., ZHANG, Q. J., SUN, Y. S., ZHENG, W. W. & CHEN, L. P. 2014. Association between dietary intake of folate and MTHFR and MTR genotype with risk of breast cancer. *Genet Mol Res,* 13**,** 8925-31.

HE, X. X., CHEN, K., YANG, J., LI, X. Y., GAN, H. Y., LIU, C. Y., COLEMAN, T. R. & AL-ABED, Y. 2009. Macrophage migration inhibitory factor promotes colorectal cancer. *Mol Med,* 15**,** 1-10.

HEDGE, T. A. & MASON, I. 2008. Expression of Shisa2, a modulator of both Wnt and Fgf signaling, in the chick embryo. *Int J Dev Biol,* 52**,** 81-5.

HENDLEY, R. L., RIESER, J. C., CAVANAGH, H. D., BODNER, B. I. & WARING, G. O., 3RD 1979. Primary radiation therapy for meibomian gland carcinoma. *Am J Ophthalmol,* 87**,** 206-9.

HENRY, J. C., PARK, J. K., JIANG, J., KIM, J. H., NAGORNEY, D. M., ROBERTS, L. R., BANERJEE, S. & SCHMITTGEN, T. D. 2010. miR-199a-3p targets CD44 and reduces proliferation of CD44 positive hepatocellular carcinoma cell lines. *Biochem Biophys Res Commun,* 403**,** 120-5.

HEO, M. J., KIM, Y. M., KOO, J. H., YANG, Y. M., AN, J., LEE, S. K., LEE, S. J., KIM, K. M., PARK, J. W. & KIM, S. G. 2014. microRNA-148a dysregulation discriminates poor prognosis of hepatocellular carcinoma in association with USP4 overexpression. *Oncotarget,* 5**,** 2792-806.

HERNANDEZ-MARTIN, A., TORRELO, A., CIRIA, S., COLMENERO, I., AGUILAR, A., GRIMALT, R. & GONZALEZ-SARMIENTO, R. 2013. Ectodermal dysplasia-skin fragility syndrome: a novel mutation in the PKP1 gene. *Clin Exp Dermatol,* 38**,** 787-90.

HEWITT, H. B. 1958. Studies of the dissemination and quantitative transplantation of a lymphocytic leukaemia of CBA mice. *Br J Cancer,* 12**,** 378-401.

HIRATA, H., HINODA, Y., SHAHRYARI, V., DENG, G., NAKAJIMA, K., TABATABAI, Z. L., ISHII, N. & DAHIYA, R. 2015. Long Noncoding RNA MALAT1 Promotes Aggressive Renal Cell Carcinoma through Ezh2 and Interacts with miR-205. *Cancer Res,* 75**,** 1322-31.

HIROSE, S. 2014. Mutant GABA(A) receptor subunits in genetic (idiopathic) epilepsy. *Prog Brain Res,* 213**,** 55-85.

HONG, S. P., NAM, K. Y., SHIN, Y. J., KIM, K. W. & AHN, S. K. 2015. Discovery of 11beta-hydroxysteroid dehydrogenase type 1 inhibitor. *Bioorg Med Chem Lett,* 25**,** 3501-6.

HONN, K. V., GUO, Y., CAI, Y., LEE, M. J., DYSON, G., ZHANG, W. & TUCKER, S. C. 2016. 12-HETER1/GPR31, a high-affinity 12(S)-hydroxyeicosatetraenoic acid receptor, is significantly up-regulated in prostate cancer and plays a critical role in prostate cancer progression. *Faseb j,* 30**,** 2360-9.

HOPFNER, K. P., KARCHER, A., SHIN, D. S., CRAIG, L., ARTHUR, L. M., CARNEY, J. P. & TAINER, J. A. 2000. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell,* 101**,** 789-800.

HOU, J., WANG, J., LIN, C., FU, J., REN, J., LI, L., GUO, H., HAN, X. & LIU, J. 2014. Circulating MicroRNA Profiles Differ between Qi-Stagnation and Qi-

Deficiency in Coronary Heart Disease Patients with Blood Stasis Syndrome. *Evid Based Complement Alternat Med,* 2014**,** 926962.

HU, B., YING, X., WANG, J., PIRIYAPONGSA, J., JORDAN, I. K., SHENG, J., YU, F., ZHAO, P., LI, Y., WANG, H., NG, W. L., HU, S., WANG, X., WANG, C., ZHENG, X., LI, W., CURRAN, W. J. & WANG, Y. 2014. Identification of a tumor-suppressive human-specific microRNA within the FHIT tumor-suppressor gene. *Cancer Res,* 74**,** 2283-94.

HUANG, B., DENG, S., LOO, S. Y., DATTA, A., YAP, Y. L., YAN, B., OOI, C. H., DINH, T. D., ZHUO, J., TOCHHAWNG, L., GOPINADHAN, S., JEGADEESAN, T., TAN, P., SALTO-TELLEZ, M., YONG, W. P., SOONG, R., YEOH, K. G., GOH, Y. C., LOBIE, P. E., YANG, H., KUMAR, A. P., MACIVER, S. K., SO, J. B. & YAP, C. T. 2016a. Gelsolin-mediated activation of PI3K/Akt pathway is crucial for hepatocyte growth factor-induced cell scattering in gastric carcinoma. *Oncotarget*.

HUANG, S., CHEN, Y., WU, W., OUYANG, N., CHEN, J., LI, H., LIU, X., SU, F., LIN, L. & YAO, Y. 2013. miR-150 promotes human breast cancer growth and malignant behavior by targeting the pro-apoptotic purinergic P2X7 receptor. *PLoS One,* 8**,** e80707.

HUANG, S. W., CHANG, S. H., MU, S. W., JIANG, H. Y., WANG, S. T., KAO, J. K., HUANG, J. L., WU, C. Y., CHEN, Y. J. & SHIEH, J. J. 2016b. Imiquimod activates p53-dependent apoptosis in a human basal cell carcinoma cell line. *J Dermatol Sci,* 81**,** 182-91.

HUANG, Y. K. & YU, J. C. 2015. Circulating microRNAs and long non-coding RNAs in gastric cancer diagnosis: An update and review. *World J Gastroenterol,* 21**,** 9863-86.

HUBER, M., SIEGENTHALER, G., MIRANCEA, N., MARENHOLZ, I., NIZETIC, D., BREITKREUTZ, D., MISCHKE, D. & HOHL, D. 2005. Isolation and characterization of human repetin, a member of the fused gene family of the epidermal differentiation complex. *J Invest Dermatol,* 124**,** 998-1007.

HUI, C. C. & ANGERS, S. 2011. Gli proteins in development and disease. *Annu Rev Cell Dev Biol,* 27**,** 513-37.

HUO, Y., YANG, M., LIU, W., YANG, J., FU, X., LIU, D., LI, J., ZHANG, J., HUA, R. & SUN, Y. 2015. High expression of DDR1 is associated with the poor prognosis in Chinese patients with pancreatic ductal adenocarcinoma. *J Exp Clin Cancer Res,* 34**,** 88.

HUSAIN, A., BLUMENSCHEIN, G. & ESMAELI, B. 2008. Treatment and outcomes for metastatic sebaceous cell carcinoma of the eyelid. *Int J Dermatol,* 47**,** 276-9.

HUSSAIN, S., BENAVENTE, S. B., NASCIMENTO, E., DRAGONI, I., KUROWSKI, A., GILLICH, A., HUMPHREYS, P. & FRYE, M. 2009. The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. *J Cell Biol,* 186**,** 27-40.

HUTCHINSON, D. T. & SULLIVAN, R. 2015. Rubinstein-Taybi Syndrome. *J Hand Surg Am,* 40**,** 1711-2.

IGUCHI, Y., ISHIHARA, S., UCHIDA, Y., TAJIMA, K., MIZUTANI, T., KAWABATA, K. & HAGA, H. 2015. Filamin B Enhances the Invasiveness of Cancer Cells into 3D Collagen Matrices. *Cell Struct Funct,* 40**,** 61-7.

IKEHATA, H. & ONO, T. 2011. The mechanisms of UV mutagenesis. *J Radiat Res,* 52**,** 115-25.

IMAI, M., OHTA, R., VARELA, J. C., SONG, H. & TOMLINSON, S. 2007. Enhancement of antibody-dependent mechanisms of tumor cell lysis by a targeted activator of complement. *Cancer Res,* 67**,** 9535-41.

IMSLAND, F., FENG, C., BOIJE, H., BED'HOM, B., FILLON, V., DORSHORST, B., RUBIN, C. J., LIU, R., GAO, Y., GU, X., WANG, Y., GOURICHON, D., ZODY, M. C., ZECCHIN, W., VIEAUD, A., TIXIER-BOICHARD, M., HU, X., HALLBOOK, F., LI, N. & ANDERSSON, L. 2012. The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. *PLoS Genet,* 8**,** e1002775.

INSTITUTE, W. 2015. *PathCards :: Cell Cycle, Mitotic Pathway and related pathways* [Online]. Available: http://pathcards.genecards.org/card/cell_cycle_mitotic [Accessed].

ISHIDA, M., KOJIMA, F. & OKABE, H. 2013. Cathepsin K expression in basal cell carcinoma. *J Eur Acad Dermatol Venereol,* 27**,** e128-30.

ISHIGAMI, S., NATSUGOE, S., TOKUDA, K., NAKAJO, A., IWASHIGE, H., ARIDOME, K., HOKITA, S. & AIKOU, T. 2001. Invariant chain expression in gastric cancer. *Cancer Lett,* 168**,** 87-91.

ISOMURA, M., OYA, N., TACHIIRI, S., KANEYASU, Y., NISHIMURA, Y., AKIMOTO, T., HAREYAMA, M., SUGITA, T., MITSUHASHI, N., YAMASHITA, T., AOKI, M., SAI, H., HIROKAWA, Y., SAKATA, K., KARASAWA, K., TOMIDA, A., TSURUO, T., MIKI, Y., NODA, T. & HIRAOKA, M. 2008. IL12RB2 and ABCA1 genes are associated with susceptibility to radiation dermatitis. *Clin Cancer Res,* 14**,** 6683-9.

IZUMI, Y., MIYAMOTO, R., MORINO, H., YOSHIZAWA, A., NISHINAKA, K., UDAKA, F., KAMEYAMA, M., MARUYAMA, H. & KAWAKAMI, H. 2013. Cerebellar ataxia with SYNE1 mutation accompanying motor neuron disease. *Neurology,* 80**,** 600-1.

JAASKELAINEN, M., NIEMINEN, A., POKKYLA, R. M., KAUPPINEN, M., LIAKKA, A., HEIKINHEIMO, M., VASKIVUO, T. E., KLEFSTROM, J. & TAPANAINEN, J. S. 2010. Regulation of cell death in human fetal and adult ovaries--role of Bok and Bcl-X(L). *Mol Cell Endocrinol,* 330**,** 17-24.

JACKSON, R. B. & LITTLE, C. C. 1933. THE EXISTENCE OF NON-CHROMOSOMAL INFLUENCE IN THE INCIDENCE OF MAMMARY TUMORS IN MICE. *Science,* 78**,** 465-6.

JACKSON, T. R., KEARNS, B. G. & THEIBERT, A. B. 2000. Cytohesins and centaurins: mediators of PI 3-kinase-regulated Arf signaling. *Trends Biochem Sci,* 25**,** 489-95.

JAGAN, L. B.-F., V; LOGAN, P; QUTUB,M; AL-SHARIF, E; BURNIER,M 2014. Sebaceous adenomas of the eyelid and Muir-Torre syndrome. *Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO)*. Orlando, USA: ARVO.

JANES, P. W., SLAPE, C. I., FARNSWORTH, R. H., ATAPATTU, L., SCOTT, A. M. & VAIL, M. E. 2014. EphA3 biology and cancer. *Growth Factors,* 32**,** 176-89.

JAVELAUD, D., ALEXAKI, V. I., DENNLER, S., MOHAMMAD, K. S., GUISE, T. A. & MAUVIEL, A. 2011. TGF-beta/SMAD/GLI2 signaling axis in cancer progression and metastasis. *Cancer Res.* United States: 2011 Aacr.

JAYARAMAN, S. S., RAYHAN, D. J., HAZANY, S. & KOLODNEY, M. S. 2014. Mutational landscape of basal cell carcinomas by whole-exome sequencing. *J Invest Dermatol,* 134**,** 213-20.

JEONG, S. J., WANG, G., CHOI, B. D., HWANG, Y. H., KIM, B. H., KO, Y. M. & JEONG, M. J. 2015. Secretory Leukocyte Protease Inhibitor (SLPI) Increases Focal Adhesion in MC3T3 Osteoblast on Titanium Surface. *J Nanosci Nanotechnol,* 15**,** 200-4.

JIN, M., YANG, Z., YE, W., XU, H. & HUA, X. 2014. MicroRNA-150 predicts a favorable prognosis in patients with epithelial ovarian cancer, and inhibits cell invasion and metastasis by suppressing transcriptional repressor ZEB1. *PLoS One,* 9**,** e103965.

JOHNSON, R. L., ROTHMAN, A. L., XIE, J., GOODRICH, L. V., BARE, J. W., BONIFAS, J. M., QUINN, A. G., MYERS, R. M., COX, D. R., EPSTEIN, E. H., JR. & SCOTT, M. P. 1996. Human homolog of patched, a candidate gene for the basal cell nevus syndrome. *Science,* 272**,** 1668-71.

JONES, C. I., ZABOLOTSKAYA, M. V., KING, A. J., STEWART, H. J., HORNE, G. A., CHEVASSUT, T. J. & NEWBURY, S. F. 2012. Identification of circulating microRNAs as diagnostic biomarkers for use in multiple myeloma. *Br J Cancer,* 107**,** 1987-96.

JONES, J. E., CAUSEY, C. P., KNUCKLEY, B., SLACK-NOYES, J. L. & THOMPSON, P. R. 2009. Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Curr Opin Drug Discov Devel,* 12**,** 616-27.

JONES, P. A. & BAYLIN, S. B. 2002. The fundamental role of epigenetic events in cancer. *Nat Rev Genet,* 3**,** 415-28.

JUNG, H., LEE, K. P., PARK, S. J., PARK, J. H., JANG, Y. S., CHOI, S. Y., JUNG, J. G., JO, K., PARK, D. Y., YOON, J. H., LIM, D. S., HONG, G. R., CHOI, C., PARK, Y. K., LEE, J. W., HONG, H. J., KIM, S. & PARK, Y. W. 2008a. TMPRSS4 promotes invasion, migration and metastasis of human tumor cells by facilitating an epithelial-mesenchymal transition. *Oncogene,* 27**,** 2635-47.

JUNG, P., MENSSEN, A., MAYR, D. & HERMEKING, H. 2008b. AP4 encodes a c-MYC-inducible repressor of p21. *Proc Natl Acad Sci U S A,* 105**,** 15046-51.

KALOGEROPOULOU, M., VOULGARI, A., KOSTOUROU, V., SANDALTZOPOULOS, R., DIKSTEIN, R., DAVIDSON, I., TORA, L. & PINTZAS, A. 2010. TAF4b and Jun/activating protein-1 collaborate to regulate the expression of integrin alpha6 and cancer cell migration properties. *Mol Cancer Res,* 8**,** 554-68.

KANG, M. A., KIM, J. T., KIM, J. H., KIM, S. Y., KIM, Y. H., YEOM, Y. I., LEE, Y. & LEE, H. G. 2009. Upregulation of the cycline kinase subunit CKS2 increases cell proliferation rate in gastric cancer. *J Cancer Res Clin Oncol,* 135**,** 761-9.

KANG, S., MIN, H. J., KANG, M. S., JUNG, M. G. & KIM, S. 2013. Discovery of novel 2-hydroxydiarylamide derivatives as TMPRSS4 inhibitors. *Bioorg Med Chem Lett,* 23**,** 1748-51.

KARAGAS, M. R., ANDREW, A. S., NELSON, H. H., LI, Z., PUNSHON, T., SCHNED, A., MARSIT, C. J., MORRIS, J. S., MOORE, J. H., TYLER, A. L., GILBERT-DIAMOND, D., GUERINOT, M. L. & KELSEY, K. T. 2012. SLC39A2 and FSIP1 polymorphisms as potential modifiers of arsenic-related bladder cancer. *Hum Genet,* 131**,** 453-61.

KARVINEN, S., KOSMA, V. M., TAMMI, M. I. & TAMMI, R. 2003. Hyaluronan, CD44 and versican in epidermal keratinocyte tumours. *Br J Dermatol,* 148**,** 86-94.

KATAOKA, F., TSUDA, H., ARAO, T., NISHIMURA, S., TANAKA, H., NOMURA, H., CHIYODA, T., HIRASAWA, A., AKAHANE, T., NISHIO, H., NISHIO,

K. & AOKI, D. 2012. EGRI and FOSB gene expressions in cancer stroma are independent prognostic indicators for epithelial ovarian cancer receiving standard therapy. *Genes Chromosomes Cancer,* 51**,** 300-12.

KATAYAMA, K., NOGUCHI, K. & SUGIMOTO, Y. 2013. FBXO15 regulates P-glycoprotein/ABCB1 expression through the ubiquitin--proteasome pathway in cancer cells. *Cancer Sci,* 104**,** 694-702.

KATOH, Y. & KATOH, M. 2006. WNT antagonist, SFRP1, is Hedgehog signaling target. *Int J Mol Med,* 17**,** 171-5.

KAWAGUCHI, K., KINAMERI, A., SUZUKI, S., SENGA, S., KE, Y. & FUJII, H. 2016. The cancer-promoting gene fatty acid-binding protein 5 (FABP5) is epigenetically regulated during human prostate carcinogenesis. *Biochem J,* 473**,** 449-61.

KAZ, A. M., LUO, Y., DZIECIATKOWSKI, S., CHAK, A., WILLIS, J. E., UPTON, M. P., LEIDNER, R. S. & GRADY, W. M. 2012. Aberrantly methylated PKP1 in the progression of Barrett's esophagus to esophageal adenocarcinoma. *Genes Chromosomes Cancer,* 51**,** 384-93.

KEBEBEW, E., PENG, M., REIFF, E., DUH, Q. Y., CLARK, O. H. & MCMILLAN, A. 2005. ECM1 and TMPRSS4 are diagnostic markers of malignant thyroid neoplasms and improve the accuracy of fine needle aspiration biopsy. *Ann Surg,* 242**,** 353-61; discussion 361-3.

KHOCHBIN, S. 2001. Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene,* 271**,** 1-12.

KHOO, K. H., VERMA, C. S. & LANE, D. P. 2014. Drugging the p53 pathway: understanding the route to clinical efficacy. *Nat Rev Drug Discov,* 13**,** 217-36.

KILLELA, P. J., PIROZZI, C. J., REITMAN, Z. J., JONES, S., RASHEED, B. A., LIPP, E., FRIEDMAN, H., FRIEDMAN, A. H., HE, Y., MCLENDON, R. E., BIGNER, D. D. & YAN, H. 2014. The genetic landscape of anaplastic astrocytoma. *Oncotarget,* 5**,** 1452-7.

KIM, D., PERTEA, G., TRAPNELL, C., PIMENTEL, H., KELLEY, R. & SALZBERG, S. L. 2013a. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol,* 14**,** R36.

KIM, E. J., SAHAI, V., ABEL, E. V., GRIFFITH, K. A., GREENSON, J. K., TAKEBE, N., KHAN, G. N., BLAU, J. L., CRAIG, R., BALIS, U. G., ZALUPSKI, M. M. & SIMEONE, D. M. 2014. Pilot Clinical Trial of Hedgehog Pathway Inhibitor GDC-0449 (Vismodegib) in Combination with Gemcitabine in Patients with Metastatic Pancreatic Adenocarcinoma. *Clin Cancer Res,* 20**,** 5937-45.

KIM, H. G., HWANG, S. Y., AARONSON, S. A., MANDINOVA, A. & LEE, S. W. 2011. DDR1 receptor tyrosine kinase promotes prosurvival pathway through Notch1 activation. *J Biol Chem,* 286**,** 17672-81.

KIM, H. G., TAN, L., WEISBERG, E. L., LIU, F., CANNING, P., CHOI, H. G., EZELL, S. A., WU, H., ZHAO, Z., WANG, J., MANDINOVA, A., GRIFFIN, J. D., BULLOCK, A. N., LIU, Q., LEE, S. W. & GRAY, N. S. 2013b. Discovery of a potent and selective DDR1 receptor tyrosine kinase inhibitor. *ACS Chem Biol,* 8**,** 2145-50.

KIM, J., TANG, J. Y., GONG, R., LEE, J. J., CLEMONS, K. V., CHONG, C. R., CHANG, K. S., FERESHTEH, M., GARDNER, D., REYA, T., LIU, J. O., EPSTEIN, E. H., STEVENS, D. A. & BEACHY, P. A. 2010a. Itraconazole, a commonly used antifungal that inhibits Hedgehog pathway activity and cancer growth. *Cancer Cell,* 17**,** 388-99.

KIM, J. H., LEE, S. H., CHO, K. J., JANG, J. J., HONG, S. I. & LEE, J. H. 1993. Enhanced expression of the c-myc protooncogene in human intracranial meningiomas. *J Korean Med Sci,* 8**,** 68-72.

KIM, J. W., LEE, S., LUI, N., CHOI, H., MULVIHILL, M., FANG, L. T., KANG, H. C., KWON, Y. W., JABLONS, D. & KIM, I. J. 2012. A somatic TSHR mutation in a patient with lung adenocarcinoma with bronchioloalveolar carcinoma, coronary artery disease and severe chronic obstructive pulmonary disease. *Oncol Rep,* 28**,** 1225-30.

KIM, N., HONG, Y., KWON, D. & YOON, S. 2013c. Somatic mutaome profile in human cancer tissues. *Genomics Inform,* 11**,** 239-44.

KIM, N., KIM, J. E., CHOUNG, H. K., LEE, M. J. & KHWARG, S. I. 2013d. Expression of Shh and Wnt signaling pathway proteins in eyelid sebaceous gland carcinoma: clinicopathologic study. *Invest Ophthalmol Vis Sci,* 54**,** 370-7.

KIM, R., SCHELL, M. J., TEER, J. K., GREENAWALT, D. M., YANG, M. & YEATMAN, T. J. 2015. Co-evolution of somatic variation in primary and metastatic colorectal cancer may expand biopsy indications in the molecular era. *PLoS One,* 10**,** e0126670.

KIM, S., KANG, H. Y., NAM, E. H., CHOI, M. S., ZHAO, X. F., HONG, C. S., LEE, J. W., LEE, J. H. & PARK, Y. K. 2010b. TMPRSS4 induces invasion and epithelial-mesenchymal transition through upregulation of integrin alpha5 and its signaling pathways. *Carcinogenesis,* 31**,** 597-606.

KIM, Y. B., HAM, I. H., HUR, H. & LEE, D. 2016. Various ARID1A expression patterns and their clinical significance in gastric cancers. *Hum Pathol,* 49**,** 61-70.

KIMYAI-ASADI, A., GOLDBERG, L. H. & JIH, M. H. 2005. Accuracy of serial transverse cross-sections in detecting residual basal cell carcinoma at the surgical margins of an elliptical excision specimen. *J Am Acad Dermatol,* 53**,** 469-74.

KING, S. J., WORTH, D. C., SCALES, T. M., MONYPENNY, J., JONES, G. E. & PARSONS, M. 2011. beta1 integrins regulate fibroblast chemotaxis through control of N-WASP stability. *Embo j,* 30**,** 1705-18.

KISHIKAWA, M., KOYAMA, K., ISEKI, M., KOBUKE, T., YONEHARA, S., SODA, M., RON, E., TOKUNAGA, M., PRESTON, D. L., MABUCHI, K. & TOKUOKA, S. 2005. Histologic characteristics of skin cancer in Hiroshima and Nagasaki: background incidence and radiation effects. *Int J Cancer,* 117**,** 363-9.

KIVELA, T., ASKO-SELJAVAARA, S., PIHKALA, U., HOVI, L. & HEIKKONEN, J. 2001. Sebaceous carcinoma of the eyelid associated with retinoblastoma. *Ophthalmology,* 108**,** 1124-8.

KLEEFF, J., ISHIWATA, T., KUMBASAR, A., FRIESS, H., BUCHLER, M. W., LANDER, A. D. & KORC, M. 1998. The cell-surface heparan sulfate proteoglycan glypican-1 regulates growth factor action in pancreatic carcinoma cells and is overexpressed in human pancreatic cancer. *J Clin Invest,* 102**,** 1662-73.

KNOWLES, M. R., LEIGH, M. W., CARSON, J. L., DAVIS, S. D., DELL, S. D., FERKOL, T. W., OLIVIER, K. N., SAGEL, S. D., ROSENFELD, M., BURNS, K. A., MINNIX, S. L., ARMSTRONG, M. C., LORI, A., HAZUCHA, M. J., LOGES, N. T., OLBRICH, H., BECKER-HECK, A., SCHMIDTS, M., WERNER, C., OMRAN, H. & ZARIWALA, M. A. 2012. Mutations of

DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. *Thorax,* 67**,** 433-41.

KNUCKLEY, B., CAUSEY, C. P., JONES, J. E., BHATIA, M., DREYTON, C. J., OSBORNE, T. C., TAKAHARA, H. & THOMPSON, P. R. 2010. Substrate specificity and kinetic studies of PADs 1, 3, and 4 identify potent and selective inhibitors of protein arginine deiminase 3. *Biochemistry,* 49**,** 4852-63.

KOBOLDT, D. C., ZHANG, Q., LARSON, D. E., SHEN, D., MCLELLAN, M. D., LIN, L., MILLER, C. A., MARDIS, E. R., DING, L. & WILSON, R. K. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res,* 22**,** 568-76.

KORC, M. 2007. Pancreatic cancer-associated stroma production. *Am J Surg,* 194**,** S84-6.

KOTOULA, V., KRIKELIS, D., KARAVASILIS, V., KOLETSA, T., ELEFTHERAKI, A. G., TELEVANTOU, D., CHRISTODOULOU, C., DIMOUDIS, S., KORANTZIS, I., PECTASIDES, D., SYRIGOS, K. N., KOSMIDIS, P. A. & FOUNTZILAS, G. 2012. Expression of DNA repair and replication genes in non-small cell lung cancer (NSCLC): a role for thymidylate synthetase (TYMS). *BMC Cancer,* 12**,** 342.

KOYAMA, S., HONDA, T., HAYANO, T., SHINOZAKI, S., KUBO, K., KOBAYASHI, T. & SEKIGUCHI, M. 1994. [A case of lung metastasis from Meibomian gland carcinoma of eyelid with effective chemotherapy]. *Gan To Kagaku Ryoho,* 21**,** 2809-12.

KOZOMARA, A. & GRIFFITHS-JONES, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res,* 42**,** D68-73.

KRUDE, T., CHRISTOV, C. P., HYRIEN, O. & MARHEINEKE, K. 2009. Y RNA functions at the initiation step of mammalian chromosomal DNA replication. *J Cell Sci,* 122**,** 2836-45.

KRUGER, T. E., MILLER, A. H., GODWIN, A. K. & WANG, J. 2014. Bone sialoprotein and osteopontin in bone metastasis of osteotropic cancers. *Crit Rev Oncol Hematol,* 89**,** 330-41.

KU, W. C., CHIU, S. K., CHEN, Y. J., HUANG, H. H. & WU, W. G. 2009. Complementary quantitative proteomics reveals that transcription factor AP-4 mediates E-box-dependent complex formation for transcriptional repression of HDM2. *Mol Cell Proteomics,* 8**,** 2034-50.

KUAN, C. S., YEE, Y. H., SEE TOO, W. C. & FEW, L. L. 2014. Ets and GATA transcription factors play a critical role in PMA-mediated repression of the ckbeta promoter via the protein kinase C signaling pathway. *PLoS One,* 9**,** e113485.

KUSAMA, T., MUKAI, M., ENDO, H., ISHIKAWA, O., TATSUTA, M., NAKAMURA, H. & INOUE, M. 2006. Inactivation of Rho GTPases by p190 RhoGAP reduces human pancreatic cancer cell invasion and metastasis. *Cancer Sci,* 97**,** 848-53.

KUZEL, P., METELITSA, A. I., DOVER, D. C. & SALOPEK, T. G. 2012. Epidemiology of sebaceous carcinoma in Alberta, Canada, from 1988 to 2007. *J Cutan Med Surg,* 16**,** 417-23.

KWASNIAK, L. A. & GARCIA-ZUAZAGA, J. 2011. Basal cell carcinoma: evidence-based medicine and review of treatment modalities. *Int J Dermatol,* 50**,** 645-58.

KWONG, L., BIJLSMA, M. F. & ROELINK, H. 2014. Shh-mediated degradation of Hhip allows cell autonomous and non-cell autonomous Shh signalling. *Nat Commun,* 5**,** 4849.

LADANYI, M. 1997. The NPM/ALK gene fusion in the pathogenesis of anaplastic large cell lymphoma. *Cancer Surv,* 30**,** 59-75.

LAI, A. Y. & WADE, P. A. 2011. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat Rev Cancer,* 11**,** 588-96.

LAI, X., WOLKENHAUER, O. & VERA, J. 2012. Modeling miRNA regulation in cancer signaling systems: miR-34a regulation of the p53/Sirt1 signaling module. *Methods Mol Biol,* 880**,** 87-108.

LAIMER, M., LANSCHUETZER, C. M., DIEM, A. & BAUER, J. W. 2010. Herlitz junctional epidermolysis bullosa. *Dermatol Clin,* 28**,** 55-60.

LAN, Y., ZHANG, Y., WANG, J., LIN, C., ITTMANN, M. M. & WANG, F. 2008. Aberrant expression of Cks1 and Cks2 contributes to prostate tumorigenesis by promoting proliferation and inhibiting programmed cell death. *Int J Cancer,* 123**,** 543-51.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LANZARDO, S., CURCIO, C., FORNI, G. & ANTON, I. M. 2007. A role for WASP Interacting Protein, WIP, in fibroblast adhesion, spreading and migration. *Int J Biochem Cell Biol,* 39**,** 262-74.

LAPIDOT, T., SIRARD, C., VORMOOR, J., MURDOCH, B., HOANG, T., CACERES-CORTES, J., MINDEN, M., PATERSON, B., CALIGIURI, M. A. & DICK, J. E. 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature,* 367**,** 645-8.

LAUGESEN, A. & HELIN, K. 2014. Chromatin repressive complexes in stem cells, development, and cancer. *Cell Stem Cell,* 14**,** 735-51.

LAUSS, M., KRIEGNER, A., VIERLINGER, K., VISNE, I., YILDIZ, A., DILAVEROGLU, E. & NOEHAMMER, C. 2008. Consensus genes of the

literature to predict breast cancer recurrence. *Breast Cancer Res Treat,* 110**,** 235-44.

LAW, C. W., CHEN, Y., SHI, W. & SMYTH, G. K. 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol,* 15**,** R29.

LAWRENCE, M. S., STOJANOV, P., POLAK, P., KRYUKOV, G. V., CIBULSKIS, K., SIVACHENKO, A., CARTER, S. L., STEWART, C., MERMEL, C. H., ROBERTS, S. A., KIEZUN, A., HAMMERMAN, P. S., MCKENNA, A., DRIER, Y., ZOU, L., RAMOS, A. H., PUGH, T. J., STRANSKY, N., HELMAN, E., KIM, J., SOUGNEZ, C., AMBROGIO, L., NICKERSON, E., SHEFLER, E., CORTES, M. L., AUCLAIR, D., SAKSENA, G., VOET, D., NOBLE, M., DICARA, D., LIN, P., LICHTENSTEIN, L., HEIMAN, D. I., FENNELL, T., IMIELINSKI, M., HERNANDEZ, B., HODIS, E., BACA, S., DULAK, A. M., LOHR, J., LANDAU, D. A., WU, C. J., MELENDEZ-ZAJGLA, J., HIDALGO-MIRANDA, A., KOREN, A., MCCARROLL, S. A., MORA, J., LEE, R. S., CROMPTON, B., ONOFRIO, R., PARKIN, M., WINCKLER, W., ARDLIE, K., GABRIEL, S. B., ROBERTS, C. W., BIEGEL, J. A., STEGMAIER, K., BASS, A. J., GARRAWAY, L. A., MEYERSON, M., GOLUB, T. R., GORDENIN, D. A., SUNYAEV, S., LANDER, E. S. & GETZ, G. 2013. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature,* 499**,** 214-8.

LAZAR, A., LYLE, S. & CALONJE, E. 2007. Sebaceous neoplasia and Torre–Muir syndrome. *Curr Diagn Pathol,* 13**,** 301-19.

LE BOIT PE, B. G., WEEDON D 2008. Pathology & Genetics Skin Tumours. World Health Organization Classification of Tumours. Lyon: IARC Press.

LEE, D. J., SCHONLEBEN, F., BANUCHI, V. E., QIU, W., CLOSE, L. G., ASSAAD, A. M. & SU, G. H. 2010. Multiple tumor-suppressor genes on chromosome 3p contribute to head and neck squamous cell carcinoma tumorigenesis. *Cancer Biol Ther,* 10**,** 689-93.

LEE, H. & PARK, W. J. 2014. Unsaturated fatty acids, desaturases, and human health. *J Med Food,* 17**,** 189-97.

LEE, L. H., SADOT, E., IVELJA, S., VAKIANI, E., HECHTMAN, J. F., SEVINSKY, C. J., KLIMSTRA, D. S., GINTY, F. & SHIA, J. 2016. ARID1A expression in early stage colorectal adenocarcinoma: an exploration of its prognostic significance. *Hum Pathol*.

LEE, M. J., KIM, N., CHOUNG, H. K., CHOE, J. Y., KHWARG, S. I. & KIM, J. E. 2015. Increased gene copy number of HER2 and concordant protein overexpression found in a subset of eyelid sebaceous gland carcinoma indicate HER2 as a potential therapeutic target. *J Cancer Res Clin Oncol*.

LEE, S. J., CHOI, Y. L., LEE, E. J., KIM, B. G., BAE, D. S., AHN, G. H. & LEE, J. H. 2007. Increased expression of calpain 6 in uterine sarcomas and carcinosarcomas: an immunohistochemical analysis. *Int J Gynecol Cancer,* 17**,** 248-53.

LEE, W. Y., WANG, C. J., LIN, T. Y., HSIAO, C. L. & LUO, C. W. 2013. CXCL17, an orphan chemokine, acts as a novel angiogenic and anti-inflammatory factor. *Am J Physiol Endocrinol Metab,* 304**,** E32-40.

LEGHA, W., GAILLARD, S., GASCON, E., MALAPERT, P., HOCINE, M., ALONSO, S. & MOQRICH, A. 2010. stac1 and stac2 genes define discrete and distinct subsets of dorsal root ganglia neurons. *Gene Expr Patterns,* 10**,** 368-75.

LEI, Z., XU, G., WANG, L., YANG, H., LIU, X., ZHAO, J. & ZHANG, H. T. 2014. MiR-142-3p represses TGF-beta-induced growth inhibition through repression of TGFbetaR1 in non-small cell lung cancer. *Faseb j,* 28**,** 2696-704.

LEIBOVITCH, I., HUILGOL, S. C., SELVA, D., RICHARDS, S. & PAVER, R. 2005. Basal cell carcinoma treated with Mohs surgery in Australia II. Outcome at 5-year follow-up. *J Am Acad Dermatol,* 53**,** 452-7.

LEICHT, D. T., KAUSAR, T., WANG, Z., FERRER-TORRES, D., WANG, T. D., THOMAS, D. G., LIN, J., CHANG, A. C., LIN, L. & BEER, D. G. 2014. TGM2: a cell surface marker in esophageal adenocarcinomas. *J Thorac Oncol,* 9**,** 872-81.

LEONG, H. S., CHEN, K., HU, Y., LEE, S., CORBIN, J., PAKUSCH, M., MURPHY, J. M., MAJEWSKI, I. J., SMYTH, G. K., ALEXANDER, W. S., HILTON, D. J. & BLEWITT, M. E. 2013. Epigenetic regulator Smchd1 functions as a tumor suppressor. *Cancer Res,* 73**,** 1591-9.

LERNER, M. R., BOYLE, J. A., HARDIN, J. A. & STEITZ, J. A. 1981. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science,* 211**,** 400-2.

LEVANAT, S., GORLIN, R. J., FALLET, S., JOHNSON, D. R., FANTASIA, J. E. & BALE, A. E. 1996. A two-hit model for developmental defects in Gorlin syndrome. *Nat Genet,* 12**,** 85-7.

LEVI, L., LOBO, G., DOUD, M. K., VON LINTIG, J., SEACHRIST, D., TOCHTROP, G. P. & NOY, N. 2013. Genetic ablation of the fatty acid-binding protein FABP5 suppresses HER2-induced mammary tumorigenesis. *Cancer Res,* 73**,** 4770-80.

LI, G. H., ARORA, P. D., CHEN, Y., MCCULLOCH, C. A. & LIU, P. 2012a. Multifunctional roles of gelsolin in health and diseases. *Med Res Rev,* 32**,** 999-1025.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.*

LI, H. J., SHEN, Z. B., SUN, Y. H., LIU, F. L., WANG, H. S. & CHEN, W. D. 2012b. [Association of T helper cell 1 cytokines expressions with prognosis of gastric cancer patients]. *Zhonghua Wei Chang Wai Ke Za Zhi,* 15**,** 618-21.

LI, J. & LAM, M. 2015. Registered report: the microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Elife,* 4**,** e06434.

LI, J., LI, L., LI, Z., GONG, G., CHEN, P., LIU, H., WANG, J., LIU, Y. & WU, X. 2015a. The role of miR-205 in the VEGF-mediated promotion of human ovarian cancer cell invasion. *Gynecol Oncol,* 137**,** 125-33.

LI, J., LI, P., ZHAO, W., YANG, R., CHEN, S., BAI, Y., DUN, S., CHEN, X., DU, Y., WANG, Y., ZANG, W., ZHAO, G. & ZHANG, G. 2015b. Expression of long non-coding RNA DLX6-AS1 in lung adenocarcinoma. *Cancer Cell Int,* 15**,** 48.

LI, L., YAN, J., XU, J., LIU, C. Q., ZHEN, Z. J., CHEN, H. W., JI, Y., WU, Z. P., HU, J. Y., ZHENG, L. & LAU, W. Y. 2014a. CXCL17 expression predicts poor prognosis and correlates with adverse immune infiltration in hepatocellular carcinoma. *PLoS One,* 9**,** e110064.

LI, L., ZHOU, Y., SUN, L., XING, G., TIAN, C., SUN, J., ZHANG, L. & HE, F. 2007. NuSAP is degraded by APC/C-Cdh1 and its overexpression results in mitotic arrest dependent of its microtubules' affinity. *Cell Signal,* 19**,** 2046-55.

LI, N., FU, H., TIE, Y., HU, Z., KONG, W., WU, Y. & ZHENG, X. 2009. miR-34a inhibits migration and invasion by down-regulation of c-Met expression in human hepatocellular carcinoma cells. *Cancer Lett,* 275**,** 44-53.

LI, N., TANG, A., HUANG, S., LI, Z., LI, X., SHEN, S., MA, J. & WANG, X. 2013. MiR-126 suppresses colon cancer cell proliferation and invasion via inhibiting RhoA/ROCK signaling pathway. *Mol Cell Biochem,* 380**,** 107-19.

LI, R., SHI, X., LING, F., WANG, C., LIU, J., WANG, W. & LI, M. 2015c. MiR-34a suppresses ovarian cancer proliferation and motility by targeting AXL. *Tumour Biol*.

LI, T., XUE, H., GUO, Y. & GUO, K. 2014b. CDKN3 is an independent prognostic factor and promotes ovarian carcinoma cell proliferation in ovarian cancer. *Oncol Rep,* 31**,** 1825-31.

LI, X., NI, R., CHEN, J., LIU, Z., XIAO, M., JIANG, F. & LU, C. 2011. The presence of IGHG1 in human pancreatic carcinomas is associated with immune evasion mechanisms. *Pancreas,* 40**,** 753-61.

LI, X. S., TROJER, P., MATSUMURA, T., TREISMAN, J. E. & TANESE, N. 2010. Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Mol Cell Biol,* 30**,** 1673-88.

LI, Y., YANG, D., BAI, Y., MO, X., HUANG, W., YUAN, W., YIN, Z., DENG, Y., MURASHKO, O., WANG, Y., FAN, X., ZHU, C., OCORR, K., BODMER, R. & WU, X. 2008a. ZNF418, a novel human KRAB/C2H2 zinc finger protein, suppresses MAPK signaling pathway. *Mol Cell Biochem,* 310**,** 141-51.

LI, Y. H., LIU, Y., LI, Y. D., LIU, Y. H., LI, F., JU, Q., XIE, P. L. & LI, G. C. 2012c. GABA stimulates human hepatocellular carcinoma growth through overexpressed GABAA receptor theta subunit. *World J Gastroenterol,* 18**,** 2704-11.

LI, Z., HAO, Q., LUO, J., XIONG, J., ZHANG, S., WANG, T., BAI, L., WANG, W., CHEN, M., GU, L., LV, K. & CHEN, J. 2015d. USP4 inhibits p53 and NF-kappaB through deubiquitinating and stabilizing HDAC2. *Oncogene*.

LI, Z., LIU, J., TANG, F., LIU, Y., WALDUM, H. L. & CUI, G. 2008b. Expression of non-mast cell histidine decarboxylase in tumor-associated microvessels in human esophageal squamous cell carcinomas. *Apmis,* 116**,** 1034-42.

LI, Z., WU, G., SHER, R. B., KHAVANDGAR, Z., HERMANSSON, M., COX, G. A., DOSCHAK, M. R., MURSHED, M., BEIER, F. & VANCE, D. E. 2014c. Choline kinase beta is required for normal endochondral bone formation. *Biochim Biophys Acta,* 1840**,** 2112-22.

LIANG, T., GUO, L. & LIU, C. 2012. Genome-wide analysis of mir-548 gene family reveals evolutionary and functional implications. *J Biomed Biotechnol,* 2012**,** 679563.

LIANG, Y., RIDZON, D., WONG, L. & CHEN, C. 2007. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics,* 8**,** 166.

LIM, S. Y., YUZHALIN, A. E., GORDON-WEEKS, A. N. & MUSCHEL, R. J. 2016. Tumor-infiltrating monocytes/macrophages promote tumor invasion and migration by upregulating S100A8 and S100A9 expression in cancer cells. *Oncogene*.

LIN, H., HUBER, R., SCHLESSINGER, D. & MORIN, P. J. 1999. Frequent silencing of the GPC3 gene in ovarian cancer cell lines. *Cancer Res,* 59**,** 807-10.

LIN, Q., XIONG, L. W., PAN, X. F., GEN, J. F., BAO, G. L., SHA, H. F., FENG, J. X., JI, C. Y. & CHEN, M. 2012. Expression of GPC3 protein and its significance in lung squamous cell carcinoma. *Med Oncol,* 29**,** 663-9.

LING, L., MAGUIRE, J. J. & DAVENPORT, A. P. 2013. Endothelin-2, the forgotten isoform: emerging role in the cardiovascular system, ovarian development, immunology and cancer. *Br J Pharmacol,* 168**,** 283-95.

LIPS, E. H., MICHAUT, M., HOOGSTRAAT, M., MULDER, L., BESSELINK, N. J., KOUDIJS, M. J., CUPPEN, E., VOEST, E. E., BERNARDS, R., NEDERLOF, P. M., WESSELING, J., RODENHUIS, S. & WESSELS, L. F. 2015. Next generation sequencing of triple negative breast cancer to find predictors for chemotherapy response. *Breast Cancer Res,* 17**,** 134.

LISMAN, R. D., JAKOBIEC, F. A. & SMALL, P. 1989. Sebaceous carcinoma of the eyelids. The role of adjunctive cryotherapy in the management of conjunctival pagetoid spread. *Ophthalmology,* 96**,** 1021-6.

LIU, H., SHUI, I. M., PLATZ, E. A., MUCCI, L. A. & GIOVANNUCCI, E. L. 2015a. No Association of ApoE Genotype with Risk of Prostate Cancer: A Nested Case-Control Study. *Cancer Epidemiol Biomarkers Prev,* 24**,** 1632-4.

LIU, H., ZHANG, H., WANG, X., TIAN, Q., HU, Z., PENG, C., JIANG, P., WANG, T., GUO, W., CHEN, Y., LI, X., ZHANG, P. & PEI, H. 2015b. The Deubiquitylating Enzyme USP4 Cooperates with CtIP in DNA Double-Strand Break End Resection. *Cell Rep,* 13**,** 93-107.

LIU, J., PAN, S., HSIEH, M. H., NG, N., SUN, F., WANG, T., KASIBHATLA, S., SCHULLER, A. G., LI, A. G., CHENG, D., LI, J., TOMPKINS, C., PFERDEKAMPER, A., STEFFY, A., CHENG, J., KOWAL, C., PHUNG, V., GUO, G., WANG, Y., GRAHAM, M. P., FLYNN, S., BRENNER, J. C., LI, C., VILLARROEL, M. C., SCHULTZ, P. G., WU, X., MCNAMARA, P., SELLERS, W. R., PETRUZZELLI, L., BORAL, A. L., SEIDEL, H. M., MCLAUGHLIN, M. E., CHE, J., CAREY, T. E., VANASSE, G. & HARRIS, J. L. 2013. Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974. *Proc Natl Acad Sci U S A,* 110**,** 20224-9.

LIU, M., HU, C., XU, Q., CHEN, L., MA, K., XU, N. & ZHU, H. 2015c. Methylseleninicacid activates Keap1/Nrf2 pathway via up-regulating miR-200a in human esophageal squamous cell carcinoma cells. *Biosci Rep*.

LIU, M. M. & ZACK, D. J. 2013. Alternative splicing and retinal degeneration. *Clin Genet,* 84**,** 142-9.

LIU, Q. L., ZHANG, J., LIU, X. & GAO, J. Y. 2015d. Role of growth hormone in maturation and activation of dendritic cells via miR-200a and the Keap1/Nrf2 pathway. *Cell Prolif,* 48**,** 573-81.

LIU, R. Z., GRAHAM, K., GLUBRECHT, D. D., GERMAIN, D. R., MACKEY, J. R. & GODBOUT, R. 2011a. Association of FABP5 expression with poor survival in triple-negative breast cancer: implication for retinoic acid therapy. *Am J Pathol,* 178**,** 997-1008.

LIU, S., RICHARDS, S. M., LO, K., HATTON, M., FAY, A. & SULLIVAN, D. A. 2011b. Changes in gene expression in human meibomian gland dysfunction. *Invest Ophthalmol Vis Sci,* 52**,** 2727-40.

LIU, Y., MEI, C., SUN, L., LI, X., LIU, M., WANG, L., LI, Z., YIN, P., ZHAO, C., SHI, Y., QIU, S., FAN, J. & ZHA, X. 2011c. The PI3K-Akt pathway regulates calpain 6 expression, proliferation, and apoptosis. *Cell Signal,* 23**,** 827-36.

LIU, Y., WANG, Y., SUN, X., MEI, C., WANG, L., LI, Z. & ZHA, X. 2015e. miR-449a promotes liver cancer cell apoptosis by down-regulation of Calpain6 and POU2F1. *Oncotarget*.

LOMBARDI, L., TAVANO, F., MORELLI, F., LATIANO, T. P., DI SEBASTIANO, P. & MAIELLO, E. 2013. Chemokine receptor CXCR4: role in gastrointestinal cancer. *Crit Rev Oncol Hematol,* 88**,** 696-705.

LONGLEY, D. B., HARKIN, D. P. & JOHNSTON, P. G. 2003. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer,* 3**,** 330-8.

LORENZ, R. & FUHRMANN, W. 1978. Familial basal cell nevus syndrome. *Hum Genet,* 44**,** 153-63.

LORUSSO, P. M., RUDIN, C. M., REDDY, J. C., TIBES, R., WEISS, G. J., BORAD, M. J., HANN, C. L., BRAHMER, J. R., CHANG, I., DARBONNE, W. C., GRAHAM, R. A., ZERIVITZ, K. L., LOW, J. A. & VON HOFF, D. D. 2011. Phase I trial of hedgehog pathway inhibitor vismodegib (GDC-0449) in patients with refractory, locally advanced or metastatic solid tumors. *Clin Cancer Res.* United States: 2011 Aacr.

LOSER, R. & PIETZSCH, J. 2015. Cysteine cathepsins: their role in tumor progression and recent trends in the development of imaging probes. *Front Chem,* 3**,** 37.

LU, K. K., TRCKA, D. & BENDECK, M. P. 2011. Collagen stimulates discoidin domain receptor 1-mediated migration of smooth muscle cells through Src. *Cardiovasc Pathol,* 20**,** 71-6.

LU, X., NANNENGA, B. & DONEHOWER, L. A. 2005. PPM1D dephosphorylates Chk1 and p53 and abrogates cell cycle checkpoints. *Genes Dev,* 19**,** 1162-74.

LUM, L. & BEACHY, P. A. 2004. The Hedgehog response network: sensors, switches, and routers. *Science.* United States.

LUO, P., FEI, J., ZHOU, J. & ZHANG, W. 2015. microRNA-126 suppresses PAK4 expression in ovarian cancer SKOV3 cells. *Oncol Lett,* 9**,** 2225-2229.

LUPIANEZ, D. G., KRAFT, K., HEINRICH, V., KRAWITZ, P., BRANCATI, F., KLOPOCKI, E., HORN, D., KAYSERILI, H., OPITZ, J. M., LAXOVA, R., SANTOS-SIMARRO, F., GILBERT-DUSSARDIER, B., WITTLER, L., BORSCHIWER, M., HAAS, S. A., OSTERWALDER, M., FRANKE, M., TIMMERMANN, B., HECHT, J., SPIELMANN, M., VISEL, A. & MUNDLOS, S. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell,* 161**,** 1012-25.

LUQUETTI, D. V., HING, A. V., RIEDER, M. J., NICKERSON, D. A., TURNER, E. H., SMITH, J., PARK, S. & CUNNINGHAM, M. L. 2013. "Mandibulofacial dysostosis with microcephaly" caused by EFTUD2 mutations: expanding the phenotype. *Am J Med Genet A,* 161a**,** 108-13.

LV, C., BAI, Z., LIU, Z., LUO, P. & ZHANG, J. 2015. Renal cell carcinoma risk is associated with the interactions of APOE, VHL and MTHFR gene polymorphisms. *Int J Clin Exp Pathol,* 8**,** 5781-6.

LV, M., ZHANG, X., LI, M., CHEN, Q., YE, M., LIANG, W., DING, L., CAI, H., FU, D. & LV, Z. 2013. miR-26a and its target CKS2 modulate cell growth and tumorigenesis of papillary thyroid carcinoma. *PLoS One,* 8**,** e67591.

LYONS, J. P., MUELLER, U. W., JI, H., EVERETT, C., FANG, X., HSIEH, J. C., BARTH, A. M. & MCCREA, P. D. 2004. Wnt-4 activates the canonical beta-catenin-mediated Wnt pathway and binds Frizzled-6 CRD: functional implications of Wnt/beta-catenin activity in kidney epithelial cells. *Exp Cell Res,* 298**,** 369-87.

MA, L., YOUNG, J., PRABHALA, H., PAN, E., MESTDAGH, P., MUTH, D., TERUYA-FELDSTEIN, J., REINHARDT, F., ONDER, T. T., VALASTYAN, S., WESTERMANN, F., SPELEMAN, F., VANDESOMPELE, J. & WEINBERG, R. A. 2010. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol,* 12**,** 247-56.

MA, X. Q., WANG, L. P., LUO, Q. C. & CAI, J. C. 2013. [Relationship between the expression level of miR-29c and biological behavior of gastric cancer]. *Zhonghua Zhong Liu Za Zhi,* 35**,** 325-30.

MACDERMED, D. M., KHODAREV, N. N., PITRODA, S. P., EDWARDS, D. C., PELIZZARI, C. A., HUANG, L., KUFE, D. W. & WEICHSELBAUM, R. R. 2010. MUC1-associated proliferation signature predicts outcomes in lung adenocarcinoma patients. *BMC Med Genomics,* 3**,** 16.

MADAN, V., LEAR, J. T. & SZEIMIES, R. M. 2010. Non-melanoma skin cancer. *Lancet,* 375**,** 673-85.

MAJETSCHAK, M. 2011. Extracellular ubiquitin: immune modulator and endogenous opponent of damage-associated molecular pattern molecules. *J Leukoc Biol,* 89**,** 205-19.

MAKAROVA, O. V., MAKAROV, E. M., LIU, S., VORNLOCHER, H. P. & LUHRMANN, R. 2002. Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *Embo j,* 21**,** 1148-57.

MALHOTRA, R., HUILGOL, S. C., HUYNH, N. T. & SELVA, D. 2004. The Australian Mohs database, part II: periocular basal cell carcinoma outcome at 5-year follow-up. *Ophthalmology,* 111**,** 631-6.

MAMIDI, S., HONE, S. & KIRSCHFINK, M. 2015. The complement system in cancer: Ambivalence between tumour destruction and promotion. *Immunobiology*.

MARSH, D., DICKINSON, S., NEILL, G. W., MARSHALL, J. F., HART, I. R. & THOMAS, G. J. 2008. alpha vbeta 6 Integrin promotes the invasion of morphoeic basal cell carcinoma through stromal modulation. *Cancer Res,* 68**,** 3295-303.

MARSHALL, D. C., LYMAN, S. K., MCCAULEY, S., KOVALENKO, M., SPANGLER, R., LIU, C., LEE, M., O'SULLIVAN, C., BARRY-HAMILTON, V., GHERMAZIEN, H., MIKELS-VIGDAL, A., GARCIA, C. A., JORGENSEN, B., VELAYO, A. C., WANG, R., ADAMKEWICZ, J. I. & SMITH, V. 2015. Selective Allosteric Inhibition of MMP9 Is Efficacious in Preclinical Models of Ulcerative Colitis and Colorectal Cancer. *PLoS One,* 10**,** e0127063.

MARTIN, S. T., SATO, N., DHARA, S., CHANG, R., HUSTINX, S. R., ABE, T., MAITRA, A. & GOGGINS, M. 2005. Aberrant methylation of the Human Hedgehog interacting protein (HHIP) gene in pancreatic neoplasms. *Cancer Biol Ther,* 4**,** 728-33.

MARTINCORENA, I., ROSHAN, A., GERSTUNG, M., ELLIS, P., VAN LOO, P., MCLAREN, S., WEDGE, D. C., FULLAM, A., ALEXANDROV, L. B., TUBIO, J. M., STEBBINGS, L., MENZIES, A., WIDAA, S., STRATTON, M. R., JONES, P. H. & CAMPBELL, P. J. 2015. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science.* United States: American Association for the Advancement of Science.

MARTINEZ, I., WANG, J., HOBSON, K. F., FERRIS, R. L. & KHAN, S. A. 2007. Identification of differentially expressed genes in HPV-positive and HPV-negative oropharyngeal squamous cell carcinomas. *Eur J Cancer,* 43**,** 415-32.

MARTINHO, O., LONGATTO-FILHO, A., LAMBROS, M. B., MARTINS, A., PINHEIRO, C., SILVA, A., PARDAL, F., AMORIM, J., MACKAY, A., MILANEZI, F., TAMBER, N., FENWICK, K., ASHWORTH, A., REIS-FILHO, J. S., LOPES, J. M. & REIS, R. M. 2009. Expression, mutation and copy number analysis of platelet-derived growth factor receptor A (PDGFRA) and its ligand PDGFA in gliomas. *Br J Cancer,* 101**,** 973-82.

MASINI, E., FABBRONI, V., GIANNINI, L., VANNACCI, A., MESSERINI, L., PERNA, F., CORTESINI, C. & CIANCHI, F. 2005. Histamine and histidine decarboxylase up-regulation in colorectal cancer: correlation with tumor stage. *Inflamm Res,* 54 Suppl 1**,** S80-1.

MASUDA, M., SAWA, M. & YAMADA, T. 2015. Therapeutic targets in the Wnt signaling pathway: Feasibility of targeting TNIK in colorectal cancer. *Pharmacol Ther,* 156**,** 1-9.

MATEESCU, B., BATISTA, L., CARDON, M., GRUOSSO, T., DE FERAUDY, Y., MARIANI, O., NICOLAS, A., MEYNIEL, J. P., COTTU, P., SASTRE-GARAU, X. & MECHTA-GRIGORIOU, F. 2011. miR-141 and miR-200a act on ovarian tumorigenesis by controlling oxidative stress response. *Nat Med,* 17**,** 1627-35.

MATSUDA, K., MARUYAMA, H., GUO, F., KLEEFF, J., ITAKURA, J., MATSUMOTO, Y., LANDER, A. D. & KORC, M. 2001. Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells. *Cancer Res,* 61**,** 5562-9.

MATSUKI, Y., TANIMOTO, A., HAMADA, T. & SASAGURI, Y. 2003. Histidine decarboxylase expression as a new sensitive and specific marker for small cell lung carcinoma. *Mod Pathol,* 16**,** 72-8.

MATSUMOTO, C. S., NAKATSUKA, K., MATSUO, K., YATSUKA, H. & MONZEN, Y. 1995. Sebaceous carcinoma responds to radiation therapy. *Ophthalmologica,* 209**,** 280-3.

MATSUNAGA, I. & SUGITA, M. 2012. Mycoketide: a CD1c-presented antigen with important implications in mycobacterial infection. *Clin Dev Immunol,* 2012**,** 981821.

MAU-SORENSEN, M., DITTRICH, C., DIENSTMANN, R., LASSEN, U., BUCHLER, W., MARTINIUS, H. & TABERNERO, J. 2015. A phase I trial of intravenous catumaxomab: a bispecific monoclonal antibody targeting EpCAM and the T cell coreceptor CD3. *Cancer Chemother Pharmacol,* 75**,** 1065-73.

MCFADDEN, D. G., PAPAGIANNAKOPOULOS, T., TAYLOR-WEINER, A., STEWART, C., CARTER, S. L., CIBULSKIS, K., BHUTKAR, A., MCKENNA, A., DOOLEY, A., VERNON, A., SOUGNEZ, C., MALSTROM, S., HEIMANN, M., PARK, J., CHEN, F., FARAGO, A. F., DAYTON, T., SHEFLER, E., GABRIEL, S., GETZ, G. & JACKS, T. 2014. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell,* 156**,** 1298-311.

MEDINA, P. P., CARRETERO, J., BALLESTAR, E., ANGULO, B., LOPEZ-RIOS, F., ESTELLER, M. & SANCHEZ-CESPEDES, M. 2005. Transcriptional targets of the chromatin-remodelling factor SMARCA4/BRG1 in lung cancer cells. *Hum Mol Genet,* 14**,** 973-82.

MERCER, T. R., DINGER, M. E. & MATTICK, J. S. 2009. Long non-coding RNAs: insights into functions. *Nat Rev Genet,* 10**,** 155-9.

MERTENS-WALKER, I., FERNANDINI, B. C., MAHARAJ, M. S., ROCKSTROH, A., NELSON, C. C., HERINGTON, A. C. & STEPHENSON, S. A. 2015. The tumour-promoting receptor tyrosine kinase, EphB4, regulates expression of integrin-beta8 in prostate cancer cells. *BMC Cancer,* 15**,** 164.

MICHAILIDOU, K., HALL, P., GONZALEZ-NEIRA, A., GHOUSSAINI, M., DENNIS, J., MILNE, R. L., SCHMIDT, M. K., CHANG-CLAUDE, J., BOJESEN, S. E., BOLLA, M. K., WANG, Q., DICKS, E., LEE, A., TURNBULL, C., RAHMAN, N., BREAST, OVARIAN CANCER SUSCEPTIBILITY, C., FLETCHER, O., PETO, J., GIBSON, L., DOS SANTOS SILVA, I., NEVANLINNA, H., MURANEN, T. A., AITTOMAKI, K., BLOMQVIST, C., CZENE, K., IRWANTO, A., LIU, J., WAISFISZ, Q., MEIJERS-HEIJBOER, H., ADANK, M., HEREDITARY, B., OVARIAN CANCER RESEARCH GROUP, N., VAN DER LUIJT, R. B., HEIN, R., DAHMEN, N., BECKMAN, L., MEINDL, A., SCHMUTZLER, R. K., MULLER-MYHSOK, B., LICHTNER, P., HOPPER, J. L., SOUTHEY, M. C., MAKALIC, E., SCHMIDT, D. F., UITTERLINDEN, A. G., HOFMAN, A., HUNTER, D. J., CHANOCK, S. J., VINCENT, D., BACOT, F., TESSIER, D. C., CANISIUS, S., WESSELS, L. F., HAIMAN, C. A., SHAH, M., LUBEN, R., BROWN, J., LUCCARINI, C., SCHOOF, N., HUMPHREYS, K., LI, J., NORDESTGAARD, B. G., NIELSEN, S. F., FLYGER, H., COUCH, F. J., WANG, X., VACHON, C., STEVENS, K. N., LAMBRECHTS, D., MOISSE, M., PARIDAENS, R., CHRISTIAENS, M. R., RUDOLPH, A., NICKELS, S., FLESCH-JANYS, D., JOHNSON, N., AITKEN, Z., AALTONEN, K., HEIKKINEN, T., BROEKS, A., VEER, L. J., VAN DER SCHOOT, C. E., GUENEL, P., TRUONG, T., LAURENT-PUIG, P., MENEGAUX, F., MARME, F., SCHNEEWEISS, A., SOHN, C., BURWINKEL, B., ZAMORA, M. P., PEREZ, J. I., PITA, G., ALONSO, M. R., COX, A., BROCK, I. W., CROSS, S. S., REED, M. W., SAWYER, E. J., et al. 2013. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet,* 45**,** 353-61, 361e1-2.

MIELE, L., MIAO, H. & NICKOLOFF, B. J. 2006. NOTCH signaling as a novel cancer therapeutic target. *Curr Cancer Drug Targets,* 6**,** 313-23.

MINNA, E., ROMEO, P., DE CECCO, L., DUGO, M., CASSINELLI, G., PILOTTI, S., DEGL'INNOCENTI, D., LANZI, C., CASALINI, P., PIEROTTI, M. A., GRECO, A. & BORRELLO, M. G. 2014. miR-199a-3p displays tumor suppressor functions in papillary thyroid carcinoma. *Oncotarget,* 5**,** 2513-28.

MITSUHASHI, S. & NISHINO, I. 2013. Megaconial congenital muscular dystrophy due to loss-of-function mutations in choline kinase beta. *Curr Opin Neurol,* 26**,** 536-43.

MIZUGUCHI, Y., ISSE, K., SPECHT, S., LUNZ, J. G., 3RD, CORBITT, N., TAKIZAWA, T. & DEMETRIS, A. J. 2014. Small proline rich protein 2a in benign and malignant liver disease. *Hepatology,* 59**,** 1130-43.

MIZUGUCHI, Y., SPECHT, S., LUNZ, J. G., 3RD, ISSE, K., CORBITT, N., TAKIZAWA, T. & DEMETRIS, A. J. 2012. SPRR2A enhances p53 deacetylation through HDAC1 and down regulates p21 promoter activity. *BMC Mol Biol,* 13**,** 20.

MLAKAR, V., BERGINC, G., VOLAVSEK, M., STOR, Z., REMS, M. & GLAVAC, D. 2009. Presence of activating KRAS mutations correlates significantly with

expression of tumour suppressor genes DCN and TPM1 in colorectal cancer. *BMC Cancer,* 9**,** 282.

MOK, T. S., WU, Y. L., THONGPRASERT, S., YANG, C. H., CHU, D. T., SAIJO, N., SUNPAWERAVONG, P., HAN, B., MARGONO, B., ICHINOSE, Y., NISHIWAKI, Y., OHE, Y., YANG, J. J., CHEWASKULYONG, B., JIANG, H., DUFFIELD, E. L., WATKINS, C. L., ARMOUR, A. A. & FUKUOKA, M. 2009. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med,* 361**,** 947-57.

MOLL, I., KURZEN, H., LANGBEIN, L. & FRANKE, W. W. 1997. The distribution of the desmosomal protein, plakophilin 1, in human skin and skin tumors. *J Invest Dermatol,* 108**,** 139-46.

MONAGHAN, P. J., LAMARCA, A., VALLE, J. W., HUBNER, R. A., MANSOOR, W., TRAINER, P. J. & DARBY, D. 2016. Routine measurement of plasma chromogranin B has limited clinical utility in the management of patients with neuroendocrine tumours. *Clin Endocrinol (Oxf),* 84**,** 348-52.

MORALES, C. R., FOX, A., EL-ALFY, M., NI, X. & ARGRAVES, W. S. 2009. Expression of Patched-1 and Smoothened in testicular meiotic and post-meiotic cells. *Microsc Res Tech,* 72**,** 809-15.

MORRIS, L. G. T. & CHAN, T. A. 2015. Therapeutic Targeting of Tumor Suppressor Genes. *Cancer,* 121**,** 1357-68.

MOSS, J. & VAUGHAN, M. 2002. Cytohesin-1 in 2001. *Arch Biochem Biophys,* 397**,** 156-61.

MOSSE, Y. P., LAUDENSLAGER, M., LONGO, L., COLE, K. A., WOOD, A., ATTIYEH, E. F., LAQUAGLIA, M. J., SENNETT, R., LYNCH, J. E., PERRI, P., LAUREYS, G., SPELEMAN, F., KIM, C., HOU, C., HAKONARSON, H., TORKAMANI, A., SCHORK, N. J., BRODEUR, G. M., TONINI, G. P., RAPPAPORT, E., DEVOTO, M. & MARIS, J. M. 2008. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature,* 455**,** 930-5.

MOSTERD, K., KREKELS, G. A., NIEMAN, F. H., OSTERTAG, J. U., ESSERS, B. A., DIRKSEN, C. D., STEIJLEN, P. M., VERMEULEN, A., NEUMANN, H. & KELLENERS-SMEETS, N. W. 2008. Surgical excision versus Mohs' micrographic surgery for primary and recurrent basal-cell carcinoma of the face: a prospective randomised controlled trial with 5-years' follow-up. *Lancet Oncol,* 9**,** 1149-56.

MOULD, A. W., PANG, Z., PAKUSCH, M., TONKS, I. D., STARK, M., CARRIE, D., MUKHOPADHYAY, P., SEIDEL, A., ELLIS, J. J., DEAKIN, J., WAKEFIELD, M. J., KRAUSE, L., BLEWITT, M. E. & KAY, G. F. 2013. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin,* 6**,** 19.

MOYA-QUILES, M. R., BERNARDO-PISA, M. V., MARTINEZ, P., GIMENO, L., BOSCH, A., SALGADO, G., MARTINEZ-BANACLOCHA, H., EGUIA, J., CAMPILLO, J. A., MURO, M., VIDAL-BUGALLO, J. B., ALVAREZ-LOPEZ, M. R. & GARCIA-ALONSO, A. M. 2013. Complement component C6 deficiency in a Spanish family: implications for clinical and molecular diagnosis. *Gene,* 521**,** 204-6.

MUIR, A. M., REN, Y., BUTZ, D. H., DAVIS, N. A., BLANK, R. D., BIRK, D. E., LEE, S. J., ROWE, D., FENG, J. Q. & GREENSPAN, D. S. 2014. Induced ablation of Bmp1 and Tll1 produces osteogenesis imperfecta in mice. *Hum Mol Genet,* 23**,** 3085-101.

MULLIGHAN, C. G., ZHANG, J., KASPER, L. H., LERACH, S., PAYNE-TURNER, D., PHILLIPS, L. A., HEATLEY, S. L., HOLMFELDT, L., COLLINS-UNDERWOOD, J. R., MA, J., BUETOW, K. H., PUI, C. H., BAKER, S. D., BRINDLE, P. K. & DOWNING, J. R. 2011. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature,* 471**,** 235-9.

MURTHY, R., HONAVAR, S. G., BURMAN, S., VEMUGANTI, G. K., NAIK, M. N. & REDDY, V. A. 2005. Neoadjuvant chemotherapy in the management of sebaceous gland carcinoma of the eyelid with regional lymph node metastasis. *Ophthal Plast Reconstr Surg,* 21**,** 307-9.

MUSANI, V., SABOL, M., CAR, D., OZRETIC, P., KALAFATIC, D., MAURAC, I., ORESKOVIC, S. & LEVANAT, S. 2013. PTCH1 gene polymorphisms in ovarian tumors: potential protective role of c.3944T allele. *Gene,* 517**,** 55-9.

MUSSNICH, P., D'ANGELO, D., LEONE, V., CROCE, C. M. & FUSCO, A. 2013. The High Mobility Group A proteins contribute to thyroid cell transformation by regulating miR-603 and miR-10b expression. *Mol Oncol,* 7**,** 531-42.

MYNENI, V. D., MELINO, G. & KAARTINEN, M. T. 2015. Transglutaminase 2--a novel inhibitor of adipogenesis. *Cell Death Dis,* 6**,** e1868.

NADENDLA, S. K., HAZAN, A., WARD, M., HARPER, L. J., MOUTASIM, K., BIANCHI, L. S., NAASE, M., GHALI, L., THOMAS, G. J., PROWSE, D. M., PHILPOTT, M. P. & NEILL, G. W. 2011. GLI1 confers profound phenotypic changes upon LNCaP prostate cancer cells that include the acquisition of a hormone independent state. *PLoS One,* 6**,** e20271.

NAGANO, T., TAKEHARA, S., TAKAHASHI, M., AIZAWA, S. & YAMAMOTO, A. 2006. Shisa2 promotes the maturation of somitic precursors and transition to the segmental fate in Xenopus embryos. *Development,* 133**,** 4643-54.

NAIR, J., JAIN, P., CHANDOLA, U., PALVE, V., VARDHAN, N. R., REDDY, R. B., KEKATPURE, V. D., SURESH, A., KURIAKOSE, M. A. & PANDA, B. 2015. Gene and miRNA expression changes in squamous cell carcinoma of larynx and hypopharynx. *Genes Cancer,* 6**,** 328-40.

NAKAMURA, F., STOSSEL, T. P. & HARTWIG, J. H. 2011. The filamins: organizers of cell structure and function. *Cell Adh Migr,* 5**,** 160-9.

NATIONAL-CANCER-PEER-REVIEW-PROGRAMME 2014. Manual for Cancer Services: Skin Measures. 1.0 ed. London: NHS Improving Quality.

NEERINCX, M., SIE, D. L., VAN DE WIEL, M. A., VAN GRIEKEN, N. C., BURGGRAAF, J. D., DEKKER, H., EIJK, P. P., YLSTRA, B., VERHOEF, C., MEIJER, G. A., BUFFART, T. E. & VERHEUL, H. M. 2015. MiR expression profiles of paired primary colorectal cancer and metastases by next-generation sequencing. *Oncogenesis,* 4**,** e170.

NEILL, G. W., HARRISON, W. J., IKRAM, M. S., WILLIAMS, T. D., BIANCHI, L. S., NADENDLA, S. K., GREEN, J. L., GHALI, L., FRISCHAUF, A. M., O'TOOLE, E. A., ABERGER, F. & PHILPOTT, M. P. 2008. GLI1 repression of ERK activity correlates with colony formation and impaired migration in human epidermal keratinocytes. *Carcinogenesis,* 29**,** 738-46.

NELSON, B. R., WU, F., LIU, Y., ANDERSON, D. M., MCANALLY, J., LIN, W., CANNON, S. C., BASSEL-DUBY, R. & OLSON, E. N. 2013. Skeletal muscle-specific T-tubule protein STAC3 mediates voltage-induced Ca2+ release and contractility. *Proc Natl Acad Sci U S A,* 110**,** 11881-6.

NETWORK, N. C. C. 2013. Basal Cell Carcinoma and Squamous Cell
Skin Cancers. *NCCN Clinical Practice Guidelines in Oncology.* 1 ed.

NEVINS, J. R. 2001. The Rb/E2F pathway and cancer. *Hum Mol Genet,* 10**,** 699-703.

NI, C. S., S; KUO, P; CHO, F;CHONG, C; ALBERT, D 1982. Sebaceous cell carcinomas of the Ocular Adnexa. *International Ophthalomology Clinics,* 22**,** 23-61.

NICULESCU, M. D., YAMAMURO, Y. & ZEISEL, S. H. 2004. Choline availability modulates human neuroblastoma cell proliferation and alters the methylation of the promoter region of the cyclin-dependent kinase inhibitor 3 gene. *J Neurochem,* 89**,** 1252-9.

NING, H., MITSUI, H., WANG, C. Q., SUAREZ-FARINAS, M., GONZALEZ, J., SHAH, K. R., CHEN, J., COATS, I., FELSEN, D., CARUCCI, J. A. & KRUEGER, J. G. 2013. Identification of anaplastic lymphoma kinase as a potential therapeutic target in Basal Cell Carcinoma. *Oncotarget,* 4**,** 2237-48.

NODA, S., ASANO, Y., TAKAHASHI, T., AKAMATA, K., AOZASA, N., TANIGUCHI, T., ICHIMURA, Y., TOYAMA, T., SUMIDA, H., KUWANO, Y., YANABA, K., TADA, Y., SUGAYA, M., KADONO, T. & SATO, S. 2013. Decreased cathepsin V expression due to Fli1 deficiency contributes to the development of dermal fibrosis and proliferative vasculopathy in systemic sclerosis. *Rheumatology (Oxford),* 52**,** 790-9.

NOGALES-GADEA, G., BRULL, A., SANTALLA, A., ANDREU, A. L., ARENAS, J., MARTIN, M. A., LUCIA, A., DE LUNA, N. & PINOS, T. 2015. McArdle Disease: Update of Reported Mutations and Polymorphisms in the PYGM Gene. *Hum Mutat,* 36**,** 669-78.

NOREAU, A., BOURASSA, C. V., SZUTO, A., LEVERT, A., DOBRZENIECKA, S., GAUTHIER, J., FORLANI, S., DURR, A., ANHEIM, M., STEVANIN, G., BRICE, A., BOUCHARD, J. P., DION, P. A., DUPRE, N. & ROULEAU, G. A. 2013. SYNE1 mutations in autosomal recessive cerebellar ataxia. *JAMA Neurol,* 70**,** 1296-31.

NOTSUDA, H., SAKURADA, A., ENDO, C., OKADA, Y., HORII, A., SHIMA, H. & KONDO, T. 2013. p190A RhoGAP is involved in EGFR pathways and promotes proliferation, invasion and migration in lung adenocarcinoma cells. *Int J Oncol,* 43**,** 1569-77.

NOTTERMAN, D. A., ALON, U., SIERK, A. J. & LEVINE, A. J. 2001. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res,* 61**,** 3124-30.

NUNERY, W. R., WELSH, M. G. & MCCORD, C. D., JR. 1983. Recurrence of sebaceous carcinoma of the eyelid after radiation therapy. *Am J Ophthalmol,* 96**,** 10-5.

O'TOOLE, J. F. & SEDOR, J. R. 2011. Are cubilin (CUBN) variants at the heart of urinary albumin excretion? *J Am Soc Nephrol.* United States.

OATLEY, J. M. & BRINSTER, R. L. 2008. Regulation of spermatogonial stem cell self-renewal in mammals. *Annu Rev Cell Dev Biol,* 24**,** 263-86.

OBATA, H., AOKI, Y., KUBOTA, S., KANAI, N. & TSURU, T. 2005. [Incidence of benign and malignant lesions of eyelid and conjunctival tumors]. *Nihon Ganka Gakkai Zasshi,* 109**,** 573-9.

OGASAWARA, M., KIM, S. C., ADAMIK, R., TOGAWA, A., FERRANS, V. J., TAKEDA, K., KIRBY, M., MOSS, J. & VAUGHAN, M. 2000. Similarities in function and gene structure of cytohesin-4 and cytohesin-1, guanine nucleotide-exchange proteins for ADP-ribosylation factors. *J Biol Chem,* 275**,** 3221-30.

OGASAWARA, S., KIYOTA, Y., CHUMAN, Y., KOWATA, A., YOSHIMURA, F., TANINO, K., KAMADA, R. & SAKAGUCHI, K. 2015. Novel inhibitors

targeting PPM1D phosphatase potently suppress cancer cell proliferation. *Bioorg Med Chem,* 23**,** 6246-9.

OGAWA, E., OWADA, Y., IKAWA, S., ADACHI, Y., EGAWA, T., NEMOTO, K., SUZUKI, K., HISHINUMA, T., KAWASHIMA, H., KONDO, H., MUTO, M., AIBA, S. & OKUYAMA, R. 2011. Epidermal FABP (FABP5) regulates keratinocyte differentiation by 13(S)-HODE-mediated activation of the NF-kappaB signaling pathway. *J Invest Dermatol,* 131**,** 604-12.

OHASHI, R., KAWAHARA, K., FUJII, T., TAKEI, H. & NAITO, Z. 2016. Higher expression of EpCAM is associated with poor clinical and pathological responses in breast cancer patients undergoing neoadjuvant chemotherapy. *Pathol Int,* 66**,** 210-7.

OHTANI, H., JIN, Z., TAKEGAWA, S., NAKAYAMA, T. & YOSHIE, O. 2009. Abundant expression of CXCL9 (MIG) by stromal cells that include dendritic cells and accumulation of CXCR3+ T cells in lymphocyte-rich gastric carcinoma. *J Pathol,* 217**,** 21-31.

OKITA, K., ICHISAKA, T. & YAMANAKA, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature,* 448**,** 313-7.

OKOSUN, J., BÖDÖR, C., WANG, J., ARAF, S., YANG, C.-Y., PAN, C., BOLLER, S., CITTARO, D., BOZEK, M., IQBAL, S., MATTHEWS, J., WRENCH, D., MARZEC, J., TAWANA, K., POPOV, N., O'RIAIN, C., O'SHEA, D., CARLOTTI, E., DAVIES, A., LAWRIE, C. H., MATOLCSY, A., CALAMINICI, M., NORTON, A., BYERS, R. J., MEIN, C., STUPKA, E., LISTER, T. A., LENZ, G., MONTOTO, S., GRIBBEN, J. G., FAN, Y., GROSSCHEDL, R., CHELALA, C. & FITZGIBBON, J. 2013. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nature Genetics,* 46**,** 176-181.

OLSEN, C. L., HSU, P. P., GLIENKE, J., RUBANYI, G. M. & BROOKS, A. R. 2004. Hedgehog-interacting protein is highly expressed in endothelial cells but down-regulated during angiogenesis and in several human tumors. *BMC Cancer,* 4**,** 43.

ORLOFF, M., PETERSON, C., HE, X., GANAPATHI, S., HEALD, B., YANG, Y. R., BEBEK, G., ROMIGH, T., SONG, J. H., WU, W., DAVID, S., CHENG, Y., MELTZER, S. J. & ENG, C. 2011. Germline mutations in MSR1, ASCC1, and CTHRC1 in patients with Barrett esophagus and esophageal adenocarcinoma. *Jama,* 306**,** 410-9.

ORR, H. A. 1995. Somatic mutation favors the evolution of diploidy. *Genetics,* 139**,** 1441-7.

OTSUKA, A., LEVESQUE, M. P., DUMMER, R. & KABASHIMA, K. 2015. Hedgehog signaling in basal cell carcinoma. *J Dermatol Sci,* 78**,** 95-100.

PAGANI, F. & BARALLE, F. E. 2004. Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet,* 5**,** 389-96.

PALOMERO, J., VEGLIANTE, M. C., RODRIGUEZ, M. L., EGUILEOR, A., CASTELLANO, G., PLANAS-RIGOL, E., JARES, P., RIBERA-CORTADA, I., CID, M. C., CAMPO, E. & AMADOR, V. 2014. SOX11 promotes tumor angiogenesis through transcriptional regulation of PDGFA in mantle cell lymphoma. *Blood,* 124**,** 2235-47.

PAN, D., LI, Y., LI, Z., WANG, Y., WANG, P. & LIANG, Y. 2012. Gli inhibitor GANT61 causes apoptosis in myeloid leukemia cells and acts in synergy with rapamycin. *Leuk Res,* 36**,** 742-8.

PAPADIA, C., LOUWAGIE, J., DEL RIO, P., GROOTECLAES, M., CORUZZI, A., MONTANA, C., NOVELLI, M., BORDI, C., DE' ANGELIS, G. L., BASSETT, P., BIGLEY, J., WARREN, B., ATKIN, W. & FORBES, A. 2014. FOXE1 and SYNE1 genes hypermethylation panel as promising biomarker in colitis-associated colorectal neoplasia. *Inflamm Bowel Dis,* 20**,** 271-7.

PARDO, F. S., WANG, C. C., ALBERT, D. & STRACHER, M. A. 1989. Sebaceous carcinoma of the ocular adnexa: radiotherapeutic management. *Int J Radiat Oncol Biol Phys,* 17**,** 643-7.

PARENT, A., ELDUQUE, X., CORNU, D., BELOT, L., LE CAER, J. P., GRANDAS, A., TOLEDANO, M. B. & D'AUTREAUX, B. 2015. Mammalian frataxin directly enhances sulfur transfer of NFS1 persulfide to both ISCU and free thiols. *Nat Commun,* 6**,** 5686.

PARK, H., STAEHLING, K., TSANG, M., APPLEBY, M. W., BRUNKOW, M. E., MARGINEANTU, D., HOCKENBERY, D. M., HABIB, T., LIGGITT, H. D., CARLSON, G. & IRITANI, B. M. 2012. Disruption of Fnip1 reveals a metabolic checkpoint controlling B lymphocyte development. *Immunity,* 36**,** 769-81.

PARK, I. H., KANG, J. H., LEE, K. S., NAM, S., RO, J. & KIM, J. H. 2014. Identification and clinical implications of circulating microRNAs for estrogen receptor-positive breast cancer. *Tumour Biol,* 35**,** 12173-80.

PARK, S. J., KIM, S. H., CHOI, H. S., RHEE, Y. & LIM, S. K. 2009. Fibroblast growth factor 2-induced cytoplasmic asparaginyl-tRNA synthetase promotes survival of osteoblasts by regulating anti-apoptotic PI3K/Akt signaling. *Bone,* 45**,** 994-1003.

PARK, Y., CHOI, Y. J., PARK, S. J., LEE, S. R., SUNG, H. J., PARK, K. H., KIM, S. J., CHOI, C. W., JUNG, K. Y. & KIM, B. S. 2011. Pretreatment serum level of 15-kDa granulysin might have a prognostic value in patients with diffuse large B cell lymphoma. *Acta Haematol,* 126**,** 79-86.

PARKER, R. L., HUNTSMAN, D. G., LESACK, D. W., CUPPLES, J. B., GRANT, D. R., AKBARI, M. & GILKS, C. B. 2002. Assessment of interlaboratory variation in the immunohistochemical determination of estrogen receptor status using a breast cancer tissue microarray. *Am J Clin Pathol,* 117**,** 723-8.

PEEDICAYIL, A., VIERKANT, R. A., HARTMANN, L. C., FRIDLEY, B. L., FREDERICKSEN, Z. S., WHITE, K. L., ELLIOTT, E. A., PHELAN, C. M., TSAI, Y. Y., BERCHUCK, A., IVERSEN, E. S., JR., COUCH, F. J., PEETHAMABARAN, P., LARSON, M. C., KALLI, K. R., KOSEL, M. L., SHRIDHAR, V., RIDER, D. N., LIEBOW, M., CUNNINGHAM, J. M., SCHILDKRAUT, J. M., SELLERS, T. A. & GOODE, E. L. 2010. Risk of ovarian cancer and inherited variants in relapse-associated genes. *PLoS One,* 5**,** e8884.

PEKARSKY, Y. & CROCE, C. M. 2015. Role of miR-15/16 in CLL. *Cell Death Differ,* 22**,** 6-11.

PENG, C., CHEN, H., WALLWIENER, M., MODUGNO, C., CUK, K., MADHAVAN, D., TRUMPP, A., HEIL, J., MARME, F., NEES, J., RIETHDORF, S., SCHOTT, S., SOHN, C., PANTEL, K., SCHNEEWEISS, A., YANG, R. & BURWINKEL, B. 2016. Plasma S100P level as a novel prognostic marker of metastatic breast cancer. *Breast Cancer Res Treat,* 157**,** 329-38.

PENSADO, A., DIAZ-CORRALES, F. J., DE LA CERDA, B., VALDES-SANCHEZ, L., DEL BOZ, A. A., RODRIGUEZ-MARTINEZ, D., GARCIA-DELGADO,

A. B., SEIJO, B., BHATTACHARYA, S. S. & SANCHEZ, A. 2016. Span poly-L-arginine nanoparticles are efficient non-viral vectors for PRPF31 gene delivery: An approach of gene therapy to treat retinitis pigmentosa. *Nanomedicine*.

PERISSINOTTO, D., IACOPETTI, P., BELLINA, I., DOLIANA, R., COLOMBATTI, A., PETTWAY, Z., BRONNER-FRASER, M., SHINOMURA, T., KIMATA, K., MORGELIN, M., LOFBERG, J. & PERRIS, R. 2000. Avian neural crest cell migration is diversely regulated by the two major hyaluronan-binding proteoglycans PG-M/versican and aggrecan. *Development,* 127**,** 2823-42.

PERSUY, M. A., SANZ, G., TROMELIN, A., THOMAS-DANGUIN, T., GIBRAT, J. F. & PAJOT-AUGY, E. 2015. Mammalian olfactory receptors: molecular mechanisms of odorant detection, 3D-modeling, and structure-activity relationships. *Prog Mol Biol Transl Sci,* 130**,** 1-36.

PESTALOZZI, B. C., PETERSON, H. F., GELBER, R. D., GOLDHIRSCH, A., GUSTERSON, B. A., TRIHIA, H., LINDTNER, J., CORTES-FUNES, H., SIMMONCINI, E., BYRNE, M. J., GOLOUH, R., RUDENSTAM, C. M., CASTIGLIONE-GERTSCH, M., ALLEGRA, C. J. & JOHNSTON, P. G. 1997. Prognostic importance of thymidylate synthase expression in early breast cancer. *J Clin Oncol,* 15**,** 1923-31.

PETERS, J. L., DUFNER-BEATTIE, J., XU, W., GEISER, J., LAHNER, B., SALT, D. E. & ANDREWS, G. K. 2007. Targeting of the mouse Slc39a2 (Zip2) gene reveals highly cell-specific patterns of expression, and unique functions in zinc, iron, and calcium homeostasis. *Genesis,* 45**,** 339-52.

PETERS, M. G., FARIAS, E., COLOMBO, L., FILMUS, J., PURICELLI, L. & BAL DE KIER JOFFE, E. 2003. Inhibition of invasion and metastasis by glypican-3 in a syngeneic breast cancer model. *Breast Cancer Res Treat,* 80**,** 221-32.

PETROVA, R. & JOYNER, A. L. 2014. Roles for Hedgehog signaling in adult organ homeostasis and repair. *Development.* England: 2014. Published by The Company of Biologists Ltd.

PFISTER, A. S., HADJIHANNAS, M. V., ROHRIG, W., SCHAMBONY, A. & BEHRENS, J. 2012a. Amer2 protein interacts with EB1 protein and adenomatous polyposis coli (APC) and controls microtubule stability and cell migration. *J Biol Chem,* 287**,** 35333-40.

PFISTER, A. S., TANNEBERGER, K., SCHAMBONY, A. & BEHRENS, J. 2012b. Amer2 protein is a novel negative regulator of Wnt/beta-catenin signaling involved in neuroectodermal patterning. *J Biol Chem,* 287**,** 1734-41.

PFLUGER, K. H., KOPPLER, H., JAQUES, G. & HAVEMANN, K. 1988. Peptide hormones in patients with acute leukaemia. *Eur J Clin Invest,* 18**,** 146-52.

PHALIPON, A. & CORTHESY, B. 2003. Novel functions of the polymeric Ig receptor: well beyond transport of immunoglobulins. *Trends Immunol,* 24**,** 55-8.

PHAROAH, P. D., TSAI, Y. Y., RAMUS, S. J., PHELAN, C. M., GOODE, E. L., LAWRENSON, K., BUCKLEY, M., FRIDLEY, B. L., TYRER, J. P., SHEN, H., WEBER, R., KAREVAN, R., LARSON, M. C., SONG, H., TESSIER, D. C., BACOT, F., VINCENT, D., CUNNINGHAM, J. M., DENNIS, J., DICKS, E., ABEN, K. K., ANTON-CULVER, H., ANTONENKOVA, N., ARMASU, S. M., BAGLIETTO, L., BANDERA, E. V., BECKMANN, M. W., BIRRER, M. J., BLOOM, G., BOGDANOVA, N., BRENTON, J. D., BRINTON, L. A., BROOKS-WILSON, A., BROWN, R., BUTZOW, R., CAMPBELL, I., CARNEY, M. E., CARVALHO, R. S., CHANG-CLAUDE, J., CHEN, Y. A., CHEN, Z., CHOW, W. H., CICEK, M. S., COETZEE, G., COOK, L. S.,

CRAMER, D. W., CYBULSKI, C., DANSONKA-MIESZKOWSKA, A., DESPIERRE, E., DOHERTY, J. A., DORK, T., DU BOIS, A., DURST, M., ECCLES, D., EDWARDS, R., EKICI, A. B., FASCHING, P. A., FENSTERMACHER, D., FLANAGAN, J., GAO, Y. T., GARCIA-CLOSAS, M., GENTRY-MAHARAJ, A., GILES, G., GJYSHI, A., GORE, M., GRONWALD, J., GUO, Q., HALLE, M. K., HARTER, P., HEIN, A., HEITZ, F., HILLEMANNS, P., HOATLIN, M., HOGDALL, E., HOGDALL, C. K., HOSONO, S., JAKUBOWSKA, A., JENSEN, A., KALLI, K. R., KARLAN, B. Y., KELEMEN, L. E., KIEMENEY, L. A., KJAER, S. K., KONECNY, G. E., KRAKSTAD, C., KUPRYJANCZYK, J., LAMBRECHTS, D., LAMBRECHTS, S., LE, N. D., LEE, N., LEE, J., LEMINEN, A., LIM, B. K., LISSOWSKA, J., LUBINSKI, J., LUNDVALL, L., LURIE, G., MASSUGER, L. F., MATSUO, K., MCGUIRE, V., et al. 2013. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat Genet,* 45**,** 362-70, 370e1-2.

PIERRY, C., PEROT, G., KARANIAN-PHILIPPE, M., NEUVILLE, A., GOMEZ-BROUCHET, A., CRESTANI, S. & COINDRE, J. M. 2015. Polypoid Laryngeal Inflammatory Myofibroblastic Tumors: Misleading Lesions: Description of Six Cases Showing ALK Overexpression. *Am J Clin Pathol,* 144**,** 511-6.

PILIA, G., HUGHES-BENZIE, R. M., MACKENZIE, A., BAYBAYAN, P., CHEN, E. Y., HUBER, R., NERI, G., CAO, A., FORABOSCO, A. & SCHLESSINGER, D. 1996. Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat Genet,* 12**,** 241-7.

PILLA, D. M., HAGAR, J. A., HALDAR, A. K., MASON, A. K., DEGRANDI, D., PFEFFER, K., ERNST, R. K., YAMAMOTO, M., MIAO, E. A. & COERS, J. 2014. Guanylate binding proteins promote caspase-11-dependent pyroptosis in response to cytoplasmic LPS. *Proc Natl Acad Sci U S A,* 111**,** 6046-51.

PISABARRO, M. T., LEUNG, B., KWONG, M., CORPUZ, R., FRANTZ, G. D., CHIANG, N., VANDLEN, R., DIEHL, L. J., SKELTON, N., KIM, H. S., EATON, D. & SCHMIDT, K. N. 2006. Cutting edge: novel human dendritic cell- and monocyte-attracting chemokine-like protein identified by fold recognition methods. *J Immunol,* 176**,** 2069-73.

PODLASEK, C. A., BARNETT, D. H., CLEMENS, J. Q., BAK, P. M. & BUSHMAN, W. 1999. Prostate development requires Sonic hedgehog expressed by the urogenital sinus epithelium. *Dev Biol,* 209**,** 28-39.

POLSTER, A., PERNI, S., BICHRAOUI, H. & BEAM, K. G. 2015. Stac adaptor proteins regulate trafficking and function of muscle and neuronal L-type Ca2+ channels. *Proc Natl Acad Sci U S A,* 112**,** 602-6.

PONTEN, F., REN, Z., NISTER, M., WESTERMARK, B. & PONTEN, J. 1994. Epithelial-stromal interactions in basal cell cancer: the PDGF system. *J Invest Dermatol,* 102**,** 304-9.

PONTISSO, P., MARTINI, A. & TURATO, C. 2014. Liver pro-oncogenic potential of SERPINB3. *Oncoscience,* 1**,** 502-3.

POPAT, S., MATAKIDOU, A. & HOULSTON, R. S. 2004. Thymidylate synthase expression and prognosis in colorectal cancer: a systematic review and meta-analysis. *J Clin Oncol,* 22**,** 529-36.

PRICA, F., RADON, T., CHENG, Y. & CRNOGORAC-JURCEVIC, T. 2016. The life and works of S100P - from conception to cancer. *Am J Cancer Res,* 6**,** 562-76.

PRICL, S., CORTELAZZI, B., DAL COL, V., MARSON, D., LAURINI, E., FERMEGLIA, M., LICITRA, L., PILOTTI, S., BOSSI, P. & PERRONE, F. 2014. Smoothened (SMO) receptor mutations dictate resistance to vismodegib in basal cell carcinoma. *Mol Oncol*.

PRICL, S., CORTELAZZI, B., DAL COL, V., MARSON, D., LAURINI, E., FERMEGLIA, M., LICITRA, L., PILOTTI, S., BOSSI, P. & PERRONE, F. 2015. Smoothened (SMO) receptor mutations dictate resistance to vismodegib in basal cell carcinoma. *Mol Oncol,* 9**,** 389-97.

PRIGIONE, I., COVONE, A. E., GIACOPELLI, F., BOCCA, P., RISSO, M., TRIPODI, G., PISTORIO, A., SOZZI, G., AIROLDI, I., RAVAZZOLO, R. & PISTOIA, V. 2016. IL12RB2 Polymorphisms correlate with risk of lung adenocarcinoma. *Immunobiology,* 221**,** 291-9.

PRUD'HOMME, G. J., GLINKA, Y., HASILO, C., PARASKEVAS, S., LI, X. & WANG, Q. 2013. GABA protects human islet cells against the deleterious effects of immunosuppressive drugs and exerts immunoinhibitory effects alone. *Transplantation,* 96**,** 616-23.

PRUD'HOMME, G. J., GLINKA, Y., UDOVYK, O., HASILO, C., PARASKEVAS, S. & WANG, Q. 2014. GABA protects pancreatic beta cells against apoptosis by increasing SIRT1 expression and activity. *Biochem Biophys Res Commun,* 452**,** 649-54.

PUGH, T. J., WEERARATNE, S. D., ARCHER, T. C., POMERANZ KRUMMEL, D. A., AUCLAIR, D., BOCHICCHIO, J., CARNEIRO, M. O., CARTER, S. L., CIBULSKIS, K., ERLICH, R. L., GREULICH, H., LAWRENCE, M. S., LENNON, N. J., MCKENNA, A., MELDRIM, J., RAMOS, A. H., ROSS, M. G., RUSS, C., SHEFLER, E., SIVACHENKO, A., SOGOLOFF, B., STOJANOV, P., TAMAYO, P., MESIROV, J. P., AMANI, V., TEIDER, N., SENGUPTA, S., FRANCOIS, J. P., NORTHCOTT, P. A., TAYLOR, M. D., YU, F., CRABTREE, G. R., KAUTZMAN, A. G., GABRIEL, S. B., GETZ, G., JAGER, N., JONES, D. T., LICHTER, P., PFISTER, S. M., ROBERTS, T. M., MEYERSON, M., POMEROY, S. L. & CHO, Y. J. 2012. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature,* 488**,** 106-10.

PUROW, B. 2015. Molecular Pathways: Targeting Diacylglycerol Kinase Alpha in Cancer. *Clin Cancer Res*.

QI, X., NG, K. T. P., LIAN, Q. Z., LIU, X. B., LI, C. X., GENG, W., LING, C. C., MA, Y. Y., YEUNG, W. H., TU, W. W., FAN, S. T., LO, C. M. & MAN, K. 2014. Clinical significance and therapeutic value of glutathione peroxidase 3 (GPx3) in hepatocellular carcinoma. *Oncotarget,* 5**,** 11103-20.

QIAN, Q., SHI, X., LEI, Z., ZHAN, L., LIU, R. Y., ZHAO, J., YANG, B., LIU, Z. & ZHANG, H. T. 2014. Methylated +58CpG site decreases DCN mRNA expression and enhances TGF-beta/Smad signaling in NSCLC cells with high metastatic potential. *Int J Oncol,* 44**,** 874-82.

QIAO, D., MEYER, K. & FRIEDL, A. 2013. Glypican 1 stimulates S phase entry and DNA replication in human glioma cells and normal astrocytes. *Mol Cell Biol,* 33**,** 4408-21.

QIU, F., SUN, R., DENG, N., GUO, T., CAO, Y., YU, Y., WANG, X., ZOU, B., ZHANG, S., JING, T., LING, T., XIE, J. & ZHANG, Q. 2015. miR-29a/b enhances cell migration and invasion in nasopharyngeal carcinoma progression by regulating SPARC and COL3A1 gene expression. *PLoS One,* 10**,** e0120969.

QU, J., YU, F., HONG, Y., GUO, Y., SUN, L., LI, X., ZHANG, J., ZHANG, H., SHI, R., CHEN, F. & LI, T. 2015. Underestimated PTCH1 mutation rate in sporadic keratocystic odontogenic tumors. *Oral Oncol,* 51**,** 40-5.

QUEIROZ, K. C., SPEK, C. A. & PEPPELENBOSCH, M. P. 2012. Targeting Hedgehog signaling and understanding refractory response to treatment with Hedgehog pathway inhibitors. *Drug Resist Updat.* Scotland: 2012 Elsevier Ltd.

RAEMAEKERS, T., RIBBECK, K., BEAUDOUIN, J., ANNAERT, W., VAN CAMP, M., STOCKMANS, I., SMETS, N., BOUILLON, R., ELLENBERG, J. & CARMELIET, G. 2003. NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization. *J Cell Biol,* 162**,** 1017-29.

RAHMAN, L., VOELLER, D., RAHMAN, M., LIPKOWITZ, S., ALLEGRA, C., BARRETT, J. C., KAYE, F. J. & ZAJAC-KAYE, M. 2004. Thymidylate synthase as an oncogene: a novel role for an essential DNA synthesis enzyme. *Cancer Cell,* 5**,** 341-51.

RAJKUMAR, T., VIJAYALAKSHMI, N., GOPAL, G., SABITHA, K., SHIRLEY, S., RAJA, U. M. & RAMAKRISHNAN, S. A. 2010. Identification and validation of genes involved in gastric tumorigenesis. *Cancer Cell Int,* 10**,** 45.

RAO, N. A., HIDAYAT, A. A., MCLEAN, I. W. & ZIMMERMAN, L. E. 1982. Sebaceous carcinomas of the ocular adnexa: A clinicopathologic study of 104 cases, with five-year follow-up data. *Hum Pathol,* 13**,** 113-22.

RAVE-FRANK, M., TEHRANY, N., KITZ, J., LEU, M., WEBER, H. E., BURFEIND, P., SCHLIEPHAKE, H., CANIS, M., BEISSBARTH, T., REICHARDT, H. M. & WOLFF, H. A. 2015. Prognostic value of CXCL12 and CXCR4 in inoperable head and neck squamous cell carcinoma. *Strahlenther Onkol*.

RAVNBAK, M. H. 2010. Objective determination of Fitzpatrick skin type. *Dan Med Bull,* 57**,** B4153.

RAYMOND, E., BRANDES, A. A., DITTRICH, C., FUMOLEAU, P., COUDERT, B., CLEMENT, P. M., FRENAY, M., RAMPLING, R., STUPP, R., KROS, J. M., HEINRICH, M. C., GORLIA, T., LACOMBE, D. & VAN DEN BENT, M. J. 2008. Phase II study of imatinib in patients with recurrent gliomas of various histologies: a European Organisation for Research and Treatment of Cancer Brain Tumor Group Study. *J Clin Oncol,* 26**,** 4659-65.

REICHERT, N., WURSTER, S., ULRICH, T., SCHMITT, K., HAUSER, S., PROBST, L., GOTZ, R., CETECI, F., MOLL, R., RAPP, U. & GAUBATZ, S. 2010. Lin9, a subunit of the mammalian DREAM complex, is essential for embryonic development, for survival of adult mice, and for tumor suppression. *Mol Cell Biol,* 30**,** 2896-908.

RENTOFT, M., LINDELL, K., TRAN, P., CHABES, A. L., BUCKLAND, R. J., WATT, D. L., MARJAVAARA, L., NILSSON, A. K., MELIN, B., TRYGG, J., JOHANSSON, E. & CHABES, A. 2016. Heterozygous colon cancer-associated mutations of SAMHD1 have functional significance. *Proc Natl Acad Sci U S A,* 113**,** 4723-8.

RIBEIRO, J. R., LOVASCO, L. A., VANDERHYDEN, B. C. & FREIMAN, R. N. 2014. Targeting TBP-Associated Factors in Ovarian Cancer. *Front Oncol,* 4**,** 45.

RIBEIRO, R., MONTEIRO, C., CATALAN, V., HU, P., CUNHA, V., RODRIGUEZ, A., GOMEZ-AMBROSI, J., FRAGA, A., PRINCIPE, P., LOBATO, C., LOBO, F., MORAIS, A., SILVA, V., SANCHES-MAGALHAES, J., OLIVEIRA, J., PINA, F., LOPES, C., MEDEIROS, R. & FRUHBECK, G. 2012. Obesity and

prostate cancer: gene expression signature of human periprostatic adipose tissue. *BMC Med,* 10**,** 108.

RICCIARDELLI, C., MAYNE, K., SYKES, P. J., RAYMOND, W. A., MCCAUL, K., MARSHALL, V. R., TILLEY, W. D., SKINNER, J. M. & HORSFALL, D. J. 1997. Elevated stromal chondroitin sulfate glycosaminoglycan predicts progression in early-stage prostate cancer. *Clin Cancer Res,* 3**,** 983-92.

RIGHI, L., BOLLITO, E., CEPPI, P., MIRABELLI, D., TAVAGLIONE, V., CHIUSA, L., PORPIGLIA, F., BRUNELLI, M., MARTIGNONI, G., TERRONE, C. & PAPOTTI, M. 2013. Prognostic role of bone sialoprotein in clear cell renal carcinoma. *Anticancer Res,* 33**,** 2679-87.

RIKER, A. I., ENKEMANN, S. A., FODSTAD, O., LIU, S., REN, S., MORRIS, C., XI, Y., HOWELL, P., METGE, B., SAMANT, R. S., SHEVDE, L. A., LI, W., ESCHRICH, S., DAUD, A., JU, J. & MATTA, J. 2008. The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. *BMC Med Genomics,* 1**,** 13.

RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* 43**,** e47.

RODRIGUEZ, D., MORRISON, C. J. & OVERALL, C. M. 2010. Matrix metalloproteinases: what do they not do? New substrates and biological roles identified by murine models and proteomics. *Biochim Biophys Acta,* 1803**,** 39-54.

RODRIGUEZ-NIETO, S. & SANCHEZ-CESPEDES, M. 2009. BRG1 and LKB1: tales of two tumor suppressor genes on chromosome 19p and lung cancer. *Carcinogenesis,* 30**,** 547-54.

ROGERS, J. 1985. Origins of repeated DNA. *Nature,* 317**,** 765-6.

ROMIGUIER, J., RANWEZ, V., DOUZERY, E. J. & GALTIER, N. 2010. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Res,* 20**,** 1001-9.

ROSSI, A., MAIONE, P., SACCO, P. C., SGAMBATO, A., CASALUCE, F., FERRARA, M. L., PALAZZOLO, G., CIARDIELLO, F. & GRIDELLI, C. 2014. ALK inhibitors and advanced non-small cell lung cancer (review). *Int J Oncol,* 45**,** 499-508.

ROTTY, J. D. & COULOMBE, P. A. 2012. A wound-induced keratin inhibits Src activity during keratinocyte migration and tissue repair. *J Cell Biol,* 197**,** 381-9.

ROWE, D. E., CARROLL, R. J. & DAY, C. L., JR. 1989a. Long-term recurrence rates in previously untreated (primary) basal cell carcinoma: implications for patient follow-up. *J Dermatol Surg Oncol,* 15**,** 315-28.

ROWE, D. E., CARROLL, R. J. & DAY, C. L., JR. 1989b. Mohs surgery is the treatment of choice for recurrent (previously treated) basal cell carcinoma. *J Dermatol Surg Oncol,* 15**,** 424-31.

RUSCHKE, K., ILLES, M., KERN, M., KLOTING, I., FASSHAUER, M., SCHON, M. R., KOSACKA, J., FITZL, G., KOVACS, P., STUMVOLL, M., BLUHER, M. & KLOTING, N. 2010. Repin1 maybe involved in the regulation of cell size and glucose transport in adipocytes. *Biochem Biophys Res Commun,* 400**,** 246-51.

RUSSO, M., RUSSO, G. L., DAGLIA, M., KASI, P. D., RAVI, S., NABAVI, S. F. & NABAVI, S. M. 2016. Understanding genistein in cancer: The "good" and the "bad" effects: A review. *Food Chem,* 196**,** 589-600.

RYU, B., KIM, D. S., DELUCA, A. M. & ALANI, R. M. 2007. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PLoS One,* 2**,** e594.

SAHIN, Z., SZCZEPNY, A., MCLAUGHLIN, E. A., MEISTRICH, M. L., ZHOU, W., USTUNEL, I. & LOVELAND, K. L. 2014. Dynamic Hedgehog signalling pathway activity in germline stem cells. *Andrology,* 2**,** 267-74.

SAITO, T., MITOMI, H., IMAMHASAN, A., HAYASHI, T., KURISAKI-ARAKAWA, A., MITANI, K., TAKAHASHI, M., KAJIYAMA, Y. & YAO, T. 2015. PTCH1 mutation is a frequent event in oesophageal basaloid squamous cell carcinoma. *Mutagenesis,* 30**,** 297-301.

SAKAMOTO, K., ARAGAKI, T., MORITA, K., KAWACHI, H., KAYAMORI, K., NAKANISHI, S., OMURA, K., MIKI, Y., OKADA, N., KATSUBE, K., TAKIZAWA, T. & YAMAGUCHI, A. 2011. Down-regulation of keratin 4 and keratin 13 expression in oral squamous cell carcinoma and epithelial dysplasia: a clue for histopathogenesis. *Histopathology,* 58**,** 531-42.

SAKANE, F. & KANOH, H. 1997. Molecules in focus: diacylglycerol kinase. *Int J Biochem Cell Biol,* 29**,** 1139-43.

SALDANHA, G., FLETCHER, A. & SLATER, D. N. 2003. Basal cell carcinoma: a dermatopathological and molecular biological update. *Br J Dermatol,* 148**,** 195-202.

SANCHEZ, P., CLEMENT, V. & RUIZ I ALTABA, A. 2005. Therapeutic targeting of the Hedgehog-GLI pathway in prostate cancer. *Cancer Res.* United States.

SANCHEZ-CARBAYO, M., SOCCI, N. D., RICHSTONE, L., CORTON, M., BEHRENDT, N., WULKFUHLE, J., BOCHNER, B., PETRICOIN, E. & CORDON-CARDO, C. 2007. Genomic and proteomic profiles reveal the association of gelsolin to TP53 status and bladder cancer progression. *Am J Pathol,* 171**,** 1650-8.

SAND, M., SKRYGAN, M., SAND, D., GEORGAS, D., GAMBICHLER, T., HAHN, S. A., ALTMEYER, P. & BECHARA, F. G. 2013. Comparative microarray analysis of microRNA expression profiles in primary cutaneous malignant melanoma, cutaneous malignant melanoma metastases, and benign melanocytic nevi. *Cell Tissue Res,* 351**,** 85-98.

SASAHIRA, T., KURIHARA, M., BHAWAL, U. K., UEDA, N., SHIMOMOTO, T., YAMAMOTO, K., KIRITA, T. & KUNIYASU, H. 2012. Downregulation of miR-126 induces angiogenesis and lymphangiogenesis by activation of VEGF-A in oral cancer. *Br J Cancer,* 107**,** 700-6.

SATO, F., TSUCHIYA, S., MELTZER, S. J. & SHIMIZU, K. 2011. MicroRNAs and epigenetics. *Febs j,* 278**,** 1598-609.

SATO, S., ITAMOCHI, H., OUMI, N., CHIBA, Y., OISHI, T., SHIMADA, M., CHIKUMI, J., NONAKA, M., KUDOH, A., KOMATSU, H., HARADA, T. & SUGIYAMA, T. 2016. Establishment and characterization of a novel ovarian clear cell carcinoma cell line, TU-OC-2, with loss of ARID1A expression. *Hum Cell*.

SAVARI, S., LIU, M., ZHANG, Y., SIME, W. & SJOLANDER, A. 2013. CysLT(1)R antagonists inhibit tumor growth in a xenograft model of colon cancer. *PLoS One,* 8**,** e73466.

SCHIFTER, S., WILLIAMS, E. D., CRAIG, R. K. & HANSEN, H. H. 1986. Calcitonin gene-related peptide and calcitonin in medullary thyroid carcinoma. *Clin Endocrinol (Oxf),* 25**,** 703-10.

SCHMITT-GRAEFF, A., KOENINGER, A., OLSCHEWSKI, M., HAXELMANS, S., NITSCHKE, R., BOCHATON-PIALLAT, M. L., LIFSCHITZ-MERCER, B., GABBIANI, G., LANGBEIN, L. & CZERNOBILSKY, B. 2007. The Ki67+ proliferation index correlates with increased cellular retinol-binding protein-1 and the coordinated loss of plakophilin-1 and desmoplakin during progression of cervical squamous lesions. *Histopathology,* 51**,** 87-97.

SCHWARTZ, R. A. & TORRE, D. P. 1995. The Muir-Torre syndrome: a 25-year retrospect. *J Am Acad Dermatol,* 33**,** 90-104.

SCHWINDINGER, W. F. & ROBISHAW, J. D. 2001. Heterotrimeric G-protein betagamma-dimers in growth and differentiation. *Oncogene,* 20**,** 1653-60.

SEKULIC, A., MIGDEN, M. R., LEWIS, K., HAINSWORTH, J. D., SOLOMON, J. A., YOO, S., ARRON, S. T., FRIEDLANDER, P. A., MARMUR, E., RUDIN, C. M., CHANG, A. L., DIRIX, L., HOU, J., YUE, H. & HAUSCHILD, A. 2015. Pivotal ERIVANCE basal cell carcinoma (BCC) study: 12-month update of efficacy and safety of vismodegib in advanced BCC. *J Am Acad Dermatol,* 72**,** 1021-6.e8.

SEKULIC, A., MIGDEN, M. R., ORO, A. E., DIRIX, L., LEWIS, K. D., HAINSWORTH, J. D., SOLOMON, J. A., YOO, S., ARRON, S. T., FRIEDLANDER, P. A., MARMUR, E., RUDIN, C. M., CHANG, A. L., LOW, J. A., MACKEY, H. M., YAUCH, R. L., GRAHAM, R. A., REDDY, J. C. & HAUSCHILD, A. 2012. Efficacy and safety of vismodegib in advanced basal-cell carcinoma. *N Engl J Med,* 366**,** 2171-9.

SENGUPTA, P. K., SMITH, E. M., KIM, K., MURNANE, M. J. & SMITH, B. D. 2003. DNA hypermethylation near the transcription start site of collagen alpha2(I) gene occurs in both cancer cell lines and primary colorectal cancers. *Cancer Res,* 63**,** 1789-97.

SHAHI, M. H., AFZAL, M., SINHA, S., EBERHART, C. G., REY, J. A., FAN, X. & CASTRESANA, J. S. 2011. Human hedgehog interacting protein expression and promoter methylation in medulloblastoma cell lines and primary tumor samples. *J Neurooncol,* 103**,** 287-96.

SHEN, C. H., CHEN, H. Y., LIN, M. S., LI, F. Y., CHANG, C. C., KUO, M. L., SETTLEMAN, J. & CHEN, R. H. 2008. Breast tumor kinase phosphorylates p190RhoGAP to regulate rho and ras and promote breast carcinoma growth, migration, and invasion. *Cancer Res,* 68**,** 7779-87.

SHEN, D. Y., FANG, Z. X., YOU, P., LIU, P. G., WANG, F., HUANG, C. L., YAO, X. B., CHEN, Z. X. & ZHANG, Z. Y. 2010a. Clinical significance and expression of cyclin kinase subunits 1 and 2 in hepatocellular carcinoma. *Liver Int,* 30**,** 119-25.

SHEN, D. Y., ZHAN, Y. H., WANG, Q. M., RUI, G. & ZHANG, Z. M. 2013a. Oncogenic potential of cyclin kinase subunit-2 in cholangiocarcinoma. *Liver Int,* 33**,** 137-48.

SHEN, Q., CICINNATI, V. R., ZHANG, X., IACOB, S., WEBER, F., SOTIROPOULOS, G. C., RADTKE, A., LU, M., PAUL, A., GERKEN, G. & BECKEBAUM, S. 2010b. Role of microRNA-199a-5p and discoidin domain receptor 1 in human hepatocellular carcinoma invasion. *Mol Cancer,* 9**,** 227.

SHEN, W. W., ZENG, Z., ZHU, W. X. & FU, G. H. 2013b. MiR-142-3p functions as a tumor suppressor by targeting CD133, ABCG2, and Lgr5 in colon cancer cells. *J Mol Med (Berl),* 91**,** 989-1000.

SHER, R. B., AOYAMA, C., HUEBSCH, K. A., JI, S., KERNER, J., YANG, Y., FRANKEL, W. N., HOPPEL, C. L., WOOD, P. A., VANCE, D. E. & COX, G.

A. 2006. A rostrocaudal muscular dystrophy caused by a defect in choline kinase beta, the first enzyme in phosphatidylcholine biosynthesis. *J Biol Chem,* 281**,** 4938-48.

SHERRY, S. T., WARD, M. H., KHOLODOV, M., BAKER, J., PHAN, L., SMIGIELSKI, E. M. & SIROTKIN, K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res,* 29**,** 308-11.

SHI, X., XU, J., WANG, J., CUI, M., GAO, Y., NIU, H. & JIN, H. 2015a. Expression analysis of apolipoprotein E and its associated genes in gastric cancer. *Oncol Lett,* 10**,** 1309-1314.

SHI, Y., QIU, M., WU, Y. & HAI, L. 2015b. MiR-548-3p functions as an anti-oncogenic regulator in breast cancer. *Biomed Pharmacother*.

SHIELDS, J. A., DEMIRCI, H., MARR, B. P., EAGLE, R. C., JR. & SHIELDS, C. L. 2004. Sebaceous carcinoma of the eyelids: personal experience with 60 cases. *Ophthalmology,* 111**,** 2151-7.

SHIELDS, J. A., DEMIRCI, H., MARR, B. P., EAGLE, R. C., JR. & SHIELDS, C. L. 2005. Sebaceous carcinoma of the ocular region: a review. *Surv Ophthalmol,* 50**,** 103-22.

SHIELDS, J. A., SHIELDS, C. L., DEMIRCI, H., HONAVAR, S. G. & SINGH, A. D. 2001. Experience with eyelid-sparing orbital exenteration: the 2000 Tullos O. Coston Lecture. *Ophthal Plast Reconstr Surg,* 17**,** 355-61.

SHIRAKURA, Y., MIZUNO, Y., WANG, L., IMAI, N., AMAIKE, C., SATO, E., ITO, M., NUKAYA, I., MINENO, J., TAKESAKO, K., IKEDA, H. & SHIKU, H. 2012. T-cell receptor gene therapy targeting melanoma-associated antigen-A4 inhibits human tumor growth in non-obese diabetic/SCID/gammacnull mice. *Cancer Sci,* 103**,** 17-25.

SIM, N. L., KUMAR, P., HU, J., HENIKOFF, S., SCHNEIDER, G. & NG, P. C. 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res,* 40**,** W452-7.

SINGH, R. R., CHO-VEGA, J. H., DAVULURI, Y., MA, S., KASBIDI, F., MILITO, C., LENNON, P. A., DRAKOS, E., MEDEIROS, L. J., LUTHRA, R. & VEGA, F. 2009. Sonic hedgehog signaling pathway is activated in ALK-positive anaplastic large cell lymphoma. *Cancer Res,* 69**,** 2550-8.

SIPRASHVILI, Z., WEBSTER, D. E., JOHNSTON, D., SHENOY, R. M., UNGEWICKELL, A. J., BHADURI, A., FLOCKHART, R., ZARNEGAR, B. J., CHE, Y., MESCHI, F., PUGLISI, J. D. & KHAVARI, P. A. 2016. The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat Genet,* 48**,** 53-8.

SIRAJ, A. K., BEG, S., JEHAN, Z., PRABHAKARAN, S., AHMED, M., A, R. H., AL-DAYEL, F., TULBAH, A., AJARIM, D. & AL-KURAYA, K. S. 2015. ALK alteration is a frequent event in aggressive breast cancers. *Breast Cancer Res,* 17**,** 127.

SIRNES, S., HONNE, H., AHMED, D., DANIELSEN, S. A., ROGNUM, T. O., MELING, G. I., LEITHE, E., RIVEDAL, E., LOTHE, R. A. & LIND, G. E. 2011. DNA methylation analyses of the connexin gene family reveal silencing of GJC1 (Connexin45) by promoter hypermethylation in colorectal cancer. *Epigenetics,* 6**,** 602-9.

SIZEMORE, G. M., SIZEMORE, S. T., SEACHRIST, D. D. & KERI, R. A. 2014. GABA(A) receptor pi (GABRP) stimulates basal-like breast cancer cell migration through activation of extracellular-regulated kinase 1/2 (ERK1/2). *J Biol Chem,* 289**,** 24102-13.

SJOBLOM, T., JONES, S., WOOD, L. D., PARSONS, D. W., LIN, J., BARBER, T. D., MANDELKER, D., LEARY, R. J., PTAK, J., SILLIMAN, N., SZABO, S., BUCKHAULTS, P., FARRELL, C., MEEH, P., MARKOWITZ, S. D., WILLIS, J., DAWSON, D., WILLSON, J. K., GAZDAR, A. F., HARTIGAN, J., WU, L., LIU, C., PARMIGIANI, G., PARK, B. H., BACHMAN, K. E., PAPADOPOULOS, N., VOGELSTEIN, B., KINZLER, K. W. & VELCULESCU, V. E. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science,* 314**,** 268-74.

SLACK, J. L., CAUSEY, C. P. & THOMPSON, P. R. 2011. Protein arginine deiminase 4: a target for an epigenetic cancer therapy. *Cell Mol Life Sci,* 68**,** 709-20.

SLATER, D. W., M 2014. Standards and datasets for reporting cancers: Dataset for the histological reporting of primary cutaneous basal cell carcinoma. 3 ed. London: RCPath.

SMITH, B. D., MAHONEY, A. P. & FELDMAN, R. S. 1983. Inverse correlation of collagen production to anchorage independence and tumorigenicity in W8- and M-cell lines. *Cancer Res,* 43**,** 4275-82.

SMITH, E. R., ZHANG, X. Y., CAPO-CHICHI, C. D., CHEN, X. & XU, X. X. 2011. Increased expression of Syne1/nesprin-1 facilitates nuclear envelope structure changes in embryonic stem cell differentiation. *Dev Dyn,* 240**,** 2245-55.

SMITH, M. M. & SANTISTEBAN, M. S. 1998. Genetic dissection of histone function. *Methods,* 15**,** 269-81.

SMYTH, M. J., KELLY, J. M., SUTTON, V. R., DAVIS, J. E., BROWNE, K. A., SAYERS, T. J. & TRAPANI, J. A. 2001. Unlocking the secrets of cytotoxic granule proteins. *J Leukoc Biol,* 70**,** 18-29.

SNOW, S. N., SAHL, W., LO, J. S., MOHS, F. E., WARNER, T., DEKKINGA, J. A. & FEYZI, J. 1994. Metastatic basal cell carcinoma. Report of five cases. *Cancer,* 73**,** 328-35.

SOBIN LH, G. M., WITTEKIND CH. 2009. International Union Against Cancer TNM Classification of Malignant Tumours. 7th Edition ed. Oxford: Wiley-Blackwell.

SONG, G. G. & LEE, Y. H. 2013. Pathway analysis of genome-wide association study on asthma. *Hum Immunol,* 74**,** 256-60.

SONG, Y. & ZUO, Y. 2014. Occurrence of HHIP gene CpG island methylation in gastric cancer. *Oncol Lett,* 8**,** 2340-2344.

SOTILLO, E., LAVER, T., MELLERT, H., SCHELTER, J. M., CLEARY, M. A., MCMAHON, S. & THOMAS-TIKHONENKO, A. 2011. Myc overexpression brings out unexpected antiapoptotic effects of miR-34a. *Oncogene,* 30**,** 2587-94.

SOUTH, A. P., WAN, H., STONE, M. G., DOPPING-HEPENSTAL, P. J., PURKIS, P. E., MARSHALL, J. F., LEIGH, I. M., EADY, R. A., HART, I. R. & MCGRATH, J. A. 2003. Lack of plakophilin 1 increases keratinocyte migration and reduces desmosome stability. *J Cell Sci,* 116**,** 3303-14.

SOUTHAM, C. M. & BRUNSCHWIG, A. 1961. Quantitative studies of autotransplantation of human cancer. Preliminary report. *Cancer,* 14**,** 971-978.

SPAGNUOLO, C., RUSSO, G. L., ORHAN, I. E., HABTEMARIAM, S., DAGLIA, M., SUREDA, A., NABAVI, S. F., DEVI, K. P., LOIZZO, M. R., TUNDIS, R. & NABAVI, S. M. 2015. Genistein and cancer: current status, challenges, and future directions. *Adv Nutr,* 6**,** 408-19.

SPECHT, S., ISSE, K., NOZAKI, I., LUNZ, J. G., 3RD & DEMETRIS, A. J. 2013. SPRR2A expression in cholangiocarcinoma increases local tumor invasiveness but prevents metastasis. *Clin Exp Metastasis,* 30**,** 877-90.

SRINIVAS, N. R. 2016. Clinical pharmacokinetics of panobinostat, a novel histone deacetylase (HDAC) inhibitor: review and perspectives. *Xenobiotica*, 1-15.

SRIVASTAVA, R. K., KAYLANI, S. Z., EDREES, N., LI, C., TALWELKAR, S. S., XU, J., PALLE, K., PRESSEY, J. G. & ATHAR, M. 2014. GLI inhibitor GANT-61 diminishes embryonal and alveolar rhabdomyosarcoma growth by inhibiting Shh/AKT-mTOR axis. *Oncotarget,* 5**,** 12151-65.

SRIVASTAVA, S. K., BHARDWAJ, A., SINGH, S., ARORA, S., WANG, B., GRIZZLE, W. E. & SINGH, A. P. 2011. MicroRNA-150 directly targets MUC4 and suppresses growth and malignant behavior of pancreatic cancer cells. *Carcinogenesis,* 32**,** 1832-9.

SROKA, I. C., ANDERSON, T. A., MCDANIEL, K. M., NAGLE, R. B., GRETZER, M. B. & CRESS, A. E. 2010. The laminin binding integrin alpha6beta1 in prostate cancer perineural invasion. *J Cell Physiol,* 224**,** 283-8.

STALEY, J. P. & WOOLFORD, J. L., JR. 2009. Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. *Curr Opin Cell Biol,* 21**,** 109-18.

STAMATOPOULOS, K., BELESSI, C., HADZIDIMITRIOU, A., SMILEVSKA, T., KALAGIAKOU, E., HATZI, K., STAVROYIANNI, N., ATHANASIADOU, A., TSOMPANAKOU, A., PAPADAKI, T., KOKKINI, G., PATERAKIS, G., SALOUM, R., LAOUTARIS, N., ANAGNOSTOPOULOS, A. & FASSAS, A. 2005. Immunoglobulin light chain repertoire in chronic lymphocytic leukemia. *Blood,* 106**,** 3575-83.

STAPLES, M. P., ELWOOD, M., BURTON, R. C., WILLIAMS, J. L., MARKS, R. & GILES, G. G. 2006. Non-melanoma skin cancer in Australia: the 2002 national survey and trends since 1985. *Med J Aust,* 184**,** 6-10.

STARK, M. S., TYAGI, S., NANCARROW, D. J., BOYLE, G. M., COOK, A. L., WHITEMAN, D. C., PARSONS, P. G., SCHMIDT, C., STURM, R. A. & HAYWARD, N. K. 2010. Characterization of the Melanoma miRNAome by Deep Sequencing. *PLoS One,* 5**,** e9685.

STEG, A., AMM, H. M., NOVAK, Z., FROST, A. R. & JOHNSON, M. R. 2010. Gli3 mediates cell survival and sensitivity to cyclopamine in pancreatic cancer. *Cancer Biol Ther,* 10**,** 893-902.

STENGER, S., ROSAT, J. P., BLOOM, B. R., KRENSKY, A. M. & MODLIN, R. L. 1999. Granulysin: a lethal weapon of cytolytic T cells. *Immunol Today,* 20**,** 390-4.

STERNBERG, S. H., REDDING, S., JINEK, M., GREENE, E. C. & DOUDNA, J. A. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature,* 507**,** 62-67.

SU, C. W., THARIN, S., JIN, Y., WIGHTMAN, B., SPECTOR, M., MEILI, D., TSUNG, N., RHINER, C., BOURIKAS, D., STOECKLI, E., GARRIGA, G., HORVITZ, H. R. & HENGARTNER, M. O. 2006. The short coiled-coil domain-containing protein UNC-69 cooperates with UNC-76 to regulate axonal outgrowth and normal presynaptic organization in Caenorhabditis elegans. *J Biol,* 5**,** 9.

SU, Y. T., GAO, C., LIU, Y., GUO, S., WANG, A., WANG, B., ERDJUMENT-BROMAGE, H., MIYAGI, M., TEMPST, P. & KAO, H. Y. 2013. Monoubiquitination of filamin B regulates vascular endothelial growth factor-mediated trafficking of histone deacetylase 7. *Mol Cell Biol,* 33**,** 1546-60.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T.

R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

SUH, H. N. & HAN, H. J. 2011. Collagen I regulates the self-renewal of mouse embryonic stem cells through alpha2beta1 integrin- and DDR1-dependent Bmi-1. *J Cell Physiol,* 226**,** 3422-32.

SUHGE D'AUBERMONT, P. C. & BENNETT, R. G. 1984. Failure of curettage and electrodesiccation for removal of basal cell carcinoma. *Arch Dermatol,* 120**,** 1456-60.

SUN, J., SONG, Y., CHEN, X., ZHAO, J., GAO, P., HUANG, X., XU, H. & WANG, Z. 2015. Novel long non-coding RNA RP11-119F7.4 as a potential biomarker for the development and progression of gastric cancer. *Oncol Lett,* 10**,** 115-120.

SURYO RAHMANTO, Y., JUNG, J. G., WU, R. C., KOBAYASHI, Y., HEAPHY, C. M., MEEKER, A. K., WANG, T. L. & SHIH, I. M. 2016. Inactivating ARID1A Tumor Suppressor Enhances TERT Transcription and Maintains Telomere Length in Cancer Cells. *J Biol Chem*.

SUZUKI, M., WATANABE, M., NAKAMARU, Y., TAKAGI, D., TAKAHASHI, H., FUKUDA, S. & HATAKEYAMA, S. 2016. TRIM39 negatively regulates the NFkappaB-mediated signaling pathway through stabilization of Cactin. *Cell Mol Life Sci,* 73**,** 1085-101.

SVAJDLER, M., BANIK, P., POLIAKOVA, K., STRAKA, L., HRIBIKOVA, Z., KINKOR, Z., KAZAKOV, D. V., SKALOVA, A. & MICHAL, M. 2015. Sebaceous carcinoma of the breast: report of four cases and review of the literature. *Pol J Pathol,* 66**,** 142-8.

SVINGEN, T., BEVERDAM, A., VERMA, P., WILHELM, D. & KOOPMAN, P. 2007. Aard is specifically up-regulated in Sertoli cells during mouse testis differentiation. *Int J Dev Biol,* 51**,** 255-8.

TABELLINI, G., BILLI, A. M., FALA, F., CAPPELLINI, A., EVAGELISTI, C., MANZOLI, L., COCCO, L. & MARTELLI, A. M. 2004. Nuclear diacylglycerol kinase-theta is activated in response to nerve growth factor stimulation of PC12 cells. *Cell Signal,* 16**,** 1263-71.

TAFLIN, H., WETTERGREN, Y., ODIN, E., CARLSSON, G. & DERWINGER, K. 2014. Folate Levels and Polymorphisms in the Genes MTHFR, MTR, and TS in Colorectal Cancer. *Clin Med Insights Oncol,* 8**,** 15-20.

TAFT, R. J., GLAZOV, E. A., LASSMANN, T., HAYASHIZAKI, Y., CARNINCI, P. & MATTICK, J. S. 2009. Small RNAs derived from snoRNAs. *Rna,* 15**,** 1233-40.

TAKANO, M., LU, Z., GOTO, T., FUSI, L., HIGHAM, J., FRANCIS, J., WITHEY, A., HARDT, J., CLOKE, B., STAVROPOULOU, A. V., ISHIHARA, O., LAM, E. W., UNTERMAN, T. G., BROSENS, J. J. & KIM, J. J. 2007. Transcriptional cross talk between the forkhead transcription factor forkhead box O1A and the progesterone receptor coordinates cell cycle regulation and differentiation in human endometrial stromal cells. *Mol Endocrinol,* 21**,** 2334-49.

TAKEDA, T., BANNO, K., OKAWA, R., YANOKURA, M., IIJIMA, M., IRIE-KUNITOMI, H., NAKAMURA, K., IIDA, M., ADACHI, M., UMENE, K., NOGAMI, Y., MASUDA, K., KOBAYASHI, Y., TOMINAGA, E. & AOKI, D. 2016. ARID1A gene mutation in ovarian and endometrial cancers (Review). *Oncol Rep,* 35**,** 607-13.

TAKEHARA, A., HOSOKAWA, M., EGUCHI, H., OHIGASHI, H., ISHIKAWA, O., NAKAMURA, Y. & NAKAGAWA, H. 2007. Gamma-aminobutyric acid (GABA) stimulates pancreatic cancer growth through overexpressing GABAA receptor pi subunit. *Cancer Res,* 67**,** 9704-12.

TANAKA, A., WEINEL, S., NAGY, N., O'DRISCOLL, M., LAI-CHEONG, J. E., KULP-SHORTEN, C. L., KNABLE, A., CARPENTER, G., FISHER, S. A., HIRAGUN, M., YANASE, Y., HIDE, M., CALLEN, J. & MCGRATH, J. A. 2012. Germline mutation in ATR in autosomal- dominant oropharyngeal cancer syndrome. *Am J Hum Genet,* 90**,** 511-7.

TANAKA, F., MATSUZAKI, S., MIMORI, K., KITA, Y., INOUE, H. & MORI, M. 2011. Clinicopathological and biological significance of CDC28 protein kinase regulatory subunit 2 overexpression in human gastric cancer. *Int J Oncol,* 39**,** 361-72.

TANESE, K., FUKUMA, M., ISHIKO, A. & SAKAMOTO, M. 2010. Endothelin-2 is upregulated in basal cell carcinoma under control of Hedgehog signaling pathway. *Biochem Biophys Res Commun,* 391**,** 486-91.

TANG, H., WANG, H., WANG, L., WANG, Q., QIN, Y., WANG, X. & LUO, S. 2014. [Expression and significance of bone sialoprotein(BSP) in esophageal squamous cell carcinoma]. *Zhonghua Zhong Liu Za Zhi,* 36**,** 602-5.

TANIMOTO, A., MATSUKI, Y., TOMITA, T., SASAGURI, T., SHIMAJIRI, S. & SASAGURI, Y. 2004. Histidine decarboxylase expression in pancreatic endocrine cells and related tumors. *Pathol Int,* 54**,** 408-12.

TAYLOR, K. J., SIMS, A. H., LIANG, L., FARATIAN, D., MUIR, M., WALKER, G., KUSKE, B., DIXON, J. M., CAMERON, D. A., HARRISON, D. J. & LANGDON, S. P. 2010. Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res,* 12**,** R39.

TAZAWA, H., TSUCHIYA, N., IZUMIYA, M. & NAKAGAMA, H. 2007. Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. *Proc Natl Acad Sci U S A,* 104**,** 15472-7.

TELFER, N. R., COLVER, G. B. & MORTON, C. A. 2008. Guidelines for the management of basal cell carcinoma. *Br J Dermatol,* 159**,** 35-48.

TEPERINO, R., AMANN, S., BAYER, M., MCGEE, S. L., LOIPETZBERGER, A., CONNOR, T., JAEGER, C., KAMMERER, B., WINTER, L., WICHE, G., DALGAARD, K., SELVARAJ, M., GASTER, M., LEE-YOUNG, R. S., FEBBRAIO, M. A., KNAUF, C., CANI, P. D., ABERGER, F., PENNINGER, J. M., POSPISILIK, J. A. & ESTERBAUER, H. 2012. Hedgehog partial agonism drives Warburg-like metabolism in muscle and brown fat. *Cell,* 151**,** 414-26.

TESSEMA, M., WILLINK, R., DO, K., YU, Y. Y., YU, W., MACHIDA, E. O., BROCK, M., VAN NESTE, L., STIDLEY, C. A., BAYLIN, S. B. & BELINSKY, S. A. 2008. Promoter methylation of genes in and around the candidate lung cancer susceptibility locus 6q23-25. *Cancer Res,* 68**,** 1707-14.

TEYE, K., OKAMOTO, K., TANAKA, Y., UMATA, T., OHNUMA, M., MOROI, M., KIMURA, H. & TSUNEOKA, M. 2008. Expression of the TAF4b gene is induced by MYC through a non-canonical, but not canonical, E-box which contributes to its specific response to MYC. *Int J Oncol,* 33**,** 1271-80.

THAYER, S. P., DI MAGLIANO, M. P., HEISER, P. W., NIELSEN, C. M., ROBERTS, D. J., LAUWERS, G. Y., QI, Y. P., GYSIN, S., FERNANDEZ-

DEL CASTILLO, C., YAJNIK, V., ANTONIU, B., MCMAHON, M., WARSHAW, A. L. & HEBROK, M. 2003. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. *Nature,* 425**,** 851-6.

THIYAGARAJAN, S., BHATIA, N., REAGAN-SHAW, S., COZMA, D., THOMAS-TIKHONENKO, A., AHMAD, N. & SPIEGELMAN, V. S. 2007. Role of GLI2 transcription factor in growth and tumorigenicity of prostate cells. *Cancer Res.* United States.

TIAN, R., XIE, X., HAN, J., LUO, C., YONG, B., PENG, H., SHEN, J. & PENG, T. 2014. miR-199a-3p negatively regulates the progression of osteosarcoma through targeting AXL. *Am J Cancer Res,* 4**,** 738-50.

TOMASETTI, C., MARCHIONNI, L., NOWAK, M. A., PARMIGIANI, G. & VOGELSTEIN, B. 2015. Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proc Natl Acad Sci U S A,* 112**,** 118-23.

TONAMI, K., HATA, S., OJIMA, K., ONO, Y., KURIHARA, Y., AMANO, T., SATO, T., KAWAMURA, Y., KURIHARA, H. & SORIMACHI, H. 2013. Calpain-6 deficiency promotes skeletal muscle development and regeneration. *PLoS Genet,* 9**,** e1003668.

TONAMI, K., KURIHARA, Y., ABURATANI, H., UCHIJIMA, Y., ASANO, T. & KURIHARA, H. 2007. Calpain 6 is involved in microtubule stabilization and cytoskeletal organization. *Mol Cell Biol,* 27**,** 2548-61.

TONEGAWA, S. 1983. Somatic generation of antibody diversity. *Nature,* 302**,** 575-81.

TORRE, L. A., BRAY, F., SIEGEL, R. L., FERLAY, J., LORTET-TIEULENT, J. & JEMAL, A. 2015. Global cancer statistics, 2012. *CA Cancer J Clin,* 65**,** 87-108.

TOSOLINI, M., KIRILOVSKY, A., MLECNIK, B., FREDRIKSEN, T., MAUGER, S., BINDEA, G., BERGER, A., BRUNEVAL, P., FRIDMAN, W. H., PAGES, F. & GALON, J. 2011. Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, th2, treg, th17) in patients with colorectal cancer. *Cancer Res,* 71**,** 1263-71.

TRAICOFF, J. L., DE MARCHIS, L., GINSBURG, B. L., ZAMORA, R. E., KHATTAR, N. H., BLANCH, V. J., PLUMMER, S., BARGO, S. A., TEMPLETON, D. J., CASEY, G. & KAETZEL, C. S. 2003. Characterization of the human polymeric immunoglobulin receptor (PIGR) 3'UTR and differential expression of PIGR mRNA during colon tumorigenesis. *J Biomed Sci,* 10**,** 792-804.

TRAN, M. N., CHOI, W., WSZOLEK, M. F., NAVAI, N., LEE, I. L., NITTI, G., WEN, S., FLORES, E. R., SIEFKER-RADTKE, A., CZERNIAK, B., DINNEY, C., BARTON, M. & MCCONKEY, D. J. 2013. The p63 protein isoform DeltaNp63alpha inhibits epithelial-mesenchymal transition in human bladder cancer cells: role of MIR-205. *J Biol Chem,* 288**,** 3275-88.

TRNSKI, D., SABOL, M., GOJEVIC, A., MARTINIC, M., OZRETIC, P., MUSANI, V., RAMIC, S. & LEVANAT, S. 2015. GSK3beta and Gli3 play a role in activation of Hedgehog-Gli pathway in human colon cancer - Targeting GSK3beta downregulates the signaling pathway and reduces cell proliferation. *Biochim Biophys Acta,* 1852**,** 2574-84.

TSAI, F. M., SHYU, R. Y., LIN, S. C., WU, C. C. & JIANG, S. Y. 2009. Induction of apoptosis by the retinoid inducible growth regulator RIG1 depends on the NC motif in HtTA cervical cancer cells. *BMC Cell Biol,* 10**,** 15.

TSAI, Y. J., WU, S. Y., HUANG, H. Y., MA, D. H., WANG, N. K., HSIAO, C. H., CHENG, C. Y. & YEH, L. K. 2015. Expression of retinoic acid-binding proteins

and retinoic acid receptors in sebaceous cell carcinoma of the eyelids. *BMC Ophthalmol,* 15**,** 142.

TU-SEKINE, B., GOLDSCHMIDT, H., PETRO, E. & RABEN, D. M. 2013. Diacylglycerol kinase theta: regulation and stability. *Adv Biol Regul,* 53**,** 118-26.

TUNNACLIFFE, A., LIU, L., MOORE, J. K., LEVERSHA, M. A., JACKSON, M. S., PAPI, L., FERGUSON-SMITH, M. A., THIESEN, H. J. & PONDER, B. A. 1993. Duplicated KOX zinc finger gene clusters flank the centromere of human chromosome 10: evidence for a pericentric inversion during primate evolution. *Nucleic Acids Res,* 21**,** 1409-17.

TURATO, C., CANNITO, S., SIMONATO, D., VILLANO, G., MORELLO, E., TERRIN, L., QUARTA, S., BIASIOLO, A., RUVOLETTO, M., MARTINI, A., FASOLATO, S., ZANUS, G., CILLO, U., GATTA, A., PAROLA, M. & PONTISSO, P. 2015. SerpinB3 and Yap Interplay Increases Myc Oncogenic Activity. *Sci Rep,* 5**,** 17701.

TURATO, C. & PONTISSO, P. 2015. SERPINB3 (serpin peptidase inhibitor, clade B (ovalbumin), member 3). *Atlas Genet Cytogenet Oncol Haematol,* 19**,** 202-209.

UBELS, J. L., GIPSON, I. K., SPURR-MICHAUD, S. J., TISDALE, A. S., VAN DYKEN, R. E. & HATTON, M. P. 2012. Gene expression in human accessory lacrimal glands of Wolfring. *Invest Ophthalmol Vis Sci,* 53**,** 6738-47.

UCKUN, F. M., STEINHERZ, P. G., SATHER, H., TRIGG, M., ARTHUR, D., TUBERGEN, D., GAYNON, P. & REAMAN, G. 1996. CD2 antigen expression on leukemic cells as a predictor of event-free survival after chemotherapy for T-lineage acute lymphoblastic leukemia: a Children's Cancer Group study. *Blood,* 88**,** 4288-95.

UHLEN, M., FAGERBERG, L., HALLSTROM, B. M., LINDSKOG, C., OKSVOLD, P., MARDINOGLU, A., SIVERTSSON, A., KAMPF, C., SJOSTEDT, E., ASPLUND, A., OLSSON, I., EDLUND, K., LUNDBERG, E., NAVANI, S., SZIGYARTO, C. A., ODEBERG, J., DJUREINOVIC, D., TAKANEN, J. O., HOBER, S., ALM, T., EDQVIST, P. H., BERLING, H., TEGEL, H., MULDER, J., ROCKBERG, J., NILSSON, P., SCHWENK, J. M., HAMSTEN, M., VON FEILITZEN, K., FORSBERG, M., PERSSON, L., JOHANSSON, F., ZWAHLEN, M., VON HEIJNE, G., NIELSEN, J. & PONTEN, F. 2015. Proteomics. Tissue-based map of the human proteome. *Science,* 347**,** 1260419.

VAIL, M. E., MURONE, C., TAN, A., HII, L., ABEBE, D., JANES, P. W., LEE, F. T., BAER, M., PALATH, V., BEBBINGTON, C., YARRANTON, G., LLERENA, C., GARIC, S., ABRAMSON, D., CARTWRIGHT, G., SCOTT, A. M. & LACKMANN, M. 2014. Targeting EphA3 inhibits cancer growth by disrupting the tumor stromal microenvironment. *Cancer Res,* 74**,** 4470-81.

VALENCIA, K., ORMAZABAL, C., ZANDUETA, C., LUIS-RAVELO, D., ANTON, I., PAJARES, M. J., AGORRETA, J., MONTUENGA, L. M., MARTINEZ-CANARIAS, S., LEITINGER, B. & LECANDA, F. 2012. Inhibition of collagen receptor discoidin domain receptor-1 (DDR1) reduces cell survival, homing, and colonization in lung cancer bone metastasis. *Clin Cancer Res,* 18**,** 969-80.

VAN ES, J. H. & CLEVERS, H. 2005. Notch and Wnt inhibitors as potential new drugs for intestinal neoplastic disease. *Trends Mol Med,* 11**,** 496-502.

VAN KEMPEN, P. M., NOORLAG, R., SWARTZ, J. E., BOVENSCHEN, N., BRAUNIUS, W. W., VERMEULEN, J. F., VAN CANN, E. M., GROLMAN, W. & WILLEMS, S. M. 2016. Oropharyngeal squamous cell carcinomas

differentially express granzyme inhibitors. *Cancer Immunol Immunother,* 65**,** 575-85.

VAN KRUIJSDIJK, R. C., VAN DER WALL, E. & VISSEREN, F. L. 2009. Obesity and cancer: the role of dysfunctional adipose tissue. *Cancer Epidemiol Biomarkers Prev,* 18**,** 2569-78.

VANHOUTTEGHEM, A. & DJIAN, P. 2006. Basonuclins 1 and 2, whose genes share a common origin, are proteins with widely different properties and functions. *Proc Natl Acad Sci U S A,* 103**,** 12423-8.

VANHOUTTEGHEM, A., MESSIAEN, S., HERVE, F., DELHOMME, B., MOISON, D., PETIT, J. M., ROUILLER-FABRE, V., LIVERA, G. & DJIAN, P. 2014. The zinc-finger protein basonuclin 2 is required for proper mitotic arrest, prevention of premature meiotic initiation and meiotic progression in mouse male germ cells. *Development,* 141**,** 4298-310.

VASSILEV, L. T., TOVAR, C., CHEN, S., KNEZEVIC, D., ZHAO, X., SUN, H., HEIMBROOK, D. C. & CHEN, L. 2006. Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc Natl Acad Sci U S A,* 103**,** 10660-5.

VELAZQUEZ-FERNANDEZ, D., LAURELL, C., GELI, J., HOOG, A., ODEBERG, J., KJELLMAN, M., LUNDEBERG, J., HAMBERGER, B., NILSSON, P. & BACKDAHL, M. 2005. Expression profiling of adrenocortical neoplasms suggests a molecular signature of malignancy. *Surgery,* 138**,** 1087-94.

VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., SMITH, H. O., YANDELL, M., EVANS, C. A., HOLT, R. A., GOCAYNE, J. D., AMANATIDES, P., BALLEW, R. M., HUSON, D. H., WORTMAN, J. R., ZHANG, Q., KODIRA, C. D., ZHENG, X. H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P. D., ZHANG, J., GABOR MIKLOS, G. L., NELSON, C., BRODER, S., CLARK, A. G., NADEAU, J., MCKUSICK, V. A., ZINDER, N., LEVINE, A. J., ROBERTS, R. J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., DI FRANCESCO, V., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A. E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T. J., HIGGINS, M. E., JI, R. R., KE, Z., KETCHUM, K. A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G. V., MILSHINA, N., MOORE, H. M., NAIK, A. K., NARAYAN, V. A., NEELAM, B., NUSSKERN, D., RUSCH, D. B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z., WANG, A., WANG, X., WANG, J., WEI, M., WIDES, R., XIAO, C., YAN, C., et al. 2001. The sequence of the human genome. *Science,* 291**,** 1304-51.

VERMEULEN, J. F., VAN HECKE, W., SPLIET, W. G., VILLACORTA HIDALGO, J., FISCH, P., BROEKHUIZEN, R. & BOVENSCHEN, N. 2016. Pediatric Primitive Neuroectodermal Tumors of the Central Nervous System Differentially Express Granzyme Inhibitors. *PLoS One,* 11**,** e0151465.

VILLARET, D. B., WANG, T., DILLON, D., XU, J., SIVAM, D., CHEEVER, M. A. & REED, S. G. 2000. Identification of genes overexpressed in head and neck

squamous cell carcinoma using a combination of complementary DNA subtraction and microarray analysis. *Laryngoscope,* 110**,** 374-81.

VON WASIELEWSKI, R., MENGEL, M., WIESE, B., RUDIGER, T., MULLER-HERMELINK, H. K. & KREIPE, H. 2002. Tissue array technology for testing interlaboratory and interobserver reproducibility of immunohistochemical estrogen receptor analysis in a large multicenter trial. *Am J Clin Pathol,* 118**,** 675-82.

WALLRAPP, C., HAHNEL, S., MULLER-PILLASCH, F., BURGHARDT, B., IWAMURA, T., RUTHENBURGER, M., LERCH, M. M., ADLER, G. & GRESS, T. M. 2000. A novel transmembrane serine protease (TMPRSS3) overexpressed in pancreatic cancer. *Cancer Res,* 60**,** 2602-6.

WANG, C. Z., YUAN, P. & LI, Y. 2015a. miR-126 regulated breast cancer cell invasion by targeting ADAM9. *Int J Clin Exp Pathol,* 8**,** 6547-53.

WANG, H. M., LU, J. H., CHEN, W. Y. & GU, A. Q. 2015b. Upregulated lncRNA-UCA1 contributes to progression of lung cancer and is closely related to clinical diagnosis as a predictive biomarker in plasma. *Int J Clin Exp Med,* 8**,** 11824-30.

WANG, I., BENDSOE, N., KLINTEBERG, C. A., ENEJDER, A. M., ANDERSSON-ENGELS, S., SVANBERG, S. & SVANBERG, K. 2001. Photodynamic therapy vs. cryosurgery of basal cell carcinomas: results of a phase III clinical trial. *Br J Dermatol,* 144**,** 832-40.

WANG, J., GAO, Y., WANG, L., LIU, X., LI, J., WANG, Z., ZHOU, J. & WANG, K. 2013a. A variant (rs932335) in the HSD11B1 gene is associated with colorectal cancer in a Chinese population. *Eur J Cancer Prev,* 22**,** 523-8.

WANG, J., HE, Q., HAN, C., GU, H., JIN, L., LI, Q., MEI, Y. & WU, M. 2012a. p53-facilitated miR-199a-3p regulates somatic cell reprogramming. *Stem Cells,* 30**,** 1405-13.

WANG, J. J., FANG, Z. X., YE, H. M., YOU, P., CAI, M. J., DUAN, H. B., WANG, F. & ZHANG, Z. Y. 2013b. Clinical significance of overexpressed cyclin-dependent kinase subunits 1 and 2 in esophageal carcinoma. *Dis Esophagus,* 26**,** 729-36.

WANG, S., AURORA, A. B., JOHNSON, B. A., QI, X., MCANALLY, J., HILL, J. A., RICHARDSON, J. A., BASSEL-DUBY, R. & OLSON, E. N. 2008. The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev Cell,* 15**,** 261-71.

WANG, S., REN, H., XU, J., YU, Y., HAN, S., QIAO, H., CHENG, S., XU, C., AN, S., JU, B., YU, C., WANG, C., WANG, T., YANG, Z., TAYLOR, E. W. & ZHAO, L. 2015c. Diminished serum repetin levels in patients with schizophrenia and bipolar disorder. *Sci Rep,* 5**,** 7977.

WANG, T., PICKARD, A. J. & GALLO, J. M. 2016a. Histone Methylation by Temozolomide; A Classic DNA Methylating Anticancer Drug. *Anticancer Res,* 36**,** 3289-99.

WANG, W., HUANG, Y., YU, Y., YANG, Y., XU, M., CHEN, X., NI, S., QIN, Q. & HUANG, X. 2016b. Fish TRIM39 regulates cell cycle progression and exerts its antiviral function against iridovirus and nodavirus. *Fish Shellfish Immunol,* 50**,** 1-10.

WANG, Y., LI, P., WANG, S., HU, J., CHEN, X. A., WU, J., FISHER, M., OSHABEN, K., ZHAO, N., GU, Y., WANG, D. & CHEN, G. 2012b. Anticancer peptidylarginine deiminase (PAD) inhibitors regulate the autophagy flux and the mammalian target of rapamycin complex 1 activity. *J Biol Chem,* 287**,** 25941-53.

WARBURG, O. P., K; NEGELEIN, E 1924. Über den Stoffwechsel der Carcinomzelle. *Biochem Zeitschr,* 152**,** 309–344.

WATANABE, K., UENO, M., KAMIYA, D., NISHIYAMA, A., MATSUMURA, M., WATAYA, T., TAKAHASHI, J. B., NISHIKAWA, S., MUGURUMA, K. & SASAI, Y. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol,* 25**,** 681-6.

WATKINS, D. N., BERMAN, D. M. & BAYLIN, S. B. 2003. Hedgehog signaling: progenitor phenotype in small-cell lung cancer. *Cell Cycle,* 2**,** 196-8.

WEEN, M. P., OEHLER, M. K. & RICCIARDELLI, C. 2011. Role of versican, hyaluronan and CD44 in ovarian cancer metastasis. *Int J Mol Sci,* 12**,** 1009-29.

WEI, W., GILBERT, N., OOI, S. L., LAWLER, J. F., OSTERTAG, E. M., KAZAZIAN, H. H., BOEKE, J. D. & MORAN, J. V. 2001. Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol,* 21**,** 1429-39.

WEINSTEIN, E. J., HEAD, R., GRIGGS, D. W., SUN, D., EVANS, R. J., SWEARINGEN, M. L., WESTLIN, M. M. & MAZZARELLA, R. 2006. VCC-1, a novel chemokine, promotes tumor growth. *Biochem Biophys Res Commun,* 350**,** 74-81.

WEISBROD, A. B., ZHANG, L., JAIN, M., BARAK, S., QUEZADO, M. M. & KEBEBEW, E. 2013. Altered PTEN, ATRX, CHGA, CHGB, and TP53 expression are associated with aggressive VHL-associated pancreatic neuroendocrine tumors. *Horm Cancer,* 4**,** 165-75.

WEN, S.-Y., LIN, Y., YU, Y.-Q., CAO, S.-J., ZHANG, R., YANG, X.-M., LI, J., ZHANG, Y.-L., WANG, Y.-H., MA, M.-Z., SUN, W.-W., LOU, X.-L., WANG, J.-H., TENG, Y.-C. & ZHANG, Z.-G. 2014. miR-506 acts as a tumor suppressor by directly targeting the hedgehog pathway transcription factor Gli3 in human cervical cancer. *Oncogene,* 34**,** 717-725.

WERNER, T. L., WADE, M. L., AGARWAL, N., BOUCHER, K., PATEL, J., LUEBKE, A. & SHARMA, S. 2015. A pilot study of JI-101, an inhibitor of VEGFR-2, PDGFR-beta, and EphB4 receptors, in combination with everolimus and as a single agent in an ovarian cancer expansion cohort. *Invest New Drugs,* 33**,** 1217-24.

WIGHT, T. N. 2002. Versican: a versatile extracellular matrix proteoglycan in cell biology. *Curr Opin Cell Biol,* 14**,** 617-23.

WIJNHOVEN, P., KONIETZNY, R., BLACKFORD, A. N., TRAVERS, J., KESSLER, B. M., NISHI, R. & JACKSON, S. P. 2015. USP4 Auto-Deubiquitylation Promotes Homologous Recombination. *Mol Cell*.

WILKINSON, S. E., FURIC, L., BUCHANAN, G., LARSSON, O., PEDERSEN, J., FRYDENBERG, M., RISBRIDGER, G. P. & TAYLOR, R. A. 2013. Hedgehog signaling is active in human prostate cancer stroma and regulates proliferation and differentiation of adjacent epithelium. *Prostate,* 73**,** 1810-23.

WILSON, B. G., HELMING, K. C., WANG, X., KIM, Y., VAZQUEZ, F., JAGANI, Z., HAHN, W. C. & ROBERTS, C. W. 2014. Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Mol Cell Biol,* 34**,** 1136-44.

WILSON, B. G. & ROBERTS, C. W. 2011. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer,* 11**,** 481-92.

WILSON, N. H. & STOECKLI, E. T. 2013. Sonic hedgehog regulates its own receptor on postcrossing commissural axons in a glypican1-dependent manner. *Neuron,* 79**,** 478-91.

WINHAM, S. J., ARMASU, S. M., CICEK, M. S., LARSON, M. C., CUNNINGHAM, J. M., KALLI, K. R., FRIDLEY, B. L. & GOODE, E. L. 2014. Genome-wide investigation of regional blood-based DNA methylation adjusted for complete blood counts implicates BNC2 in ovarian cancer. *Genet Epidemiol,* 38**,** 457-66.

WINKEL, K., ALSHEIMER, M., OLLINGER, R. & BENAVENTE, R. 2009. Protein SYCP2 provides a link between transverse filaments and lateral elements of mammalian synaptonemal complexes. *Chromosoma,* 118**,** 259-67.

WINSTON, J. S., ASCH, H. L., ZHANG, P. J., EDGE, S. B., HYLAND, A. & ASCH, B. B. 2001. Downregulation of gelsolin correlates with the progression to breast carcinoma. *Breast Cancer Res Treat,* 65**,** 11-21.

WOLF, A., KRAUSE-GRUSZCZYNSKA, M., BIRKENMEIER, O., OSTARECK-LEDERER, A., HUTTELMAIER, S. & HATZFELD, M. 2010. Plakophilin 1 stimulates translation by promoting eIF4A1 activity. *J Cell Biol,* 188**,** 463-71.

WOLIN, S. L., SIM, S. & CHEN, X. 2012. Nuclear noncoding RNA surveillance: is the end in sight? *Trends Genet,* 28**,** 306-13.

WONG, A. K., SHANAHAN, F., CHEN, Y., LIAN, L., HA, P., HENDRICKS, K., GHAFFARI, S., ILIEV, D., PENN, B., WOODLAND, A. M., SMITH, R., SALADA, G., CARILLO, A., LAITY, K., GUPTE, J., SWEDLUND, B., TAVTIGIAN, S. V., TENG, D. H. & LEES, E. 2000. BRG1, a component of the SWI-SNF complex, is mutated in multiple human tumor cell lines. *Cancer Res,* 60**,** 6171-7.

WORST, T. S., REINER, V., GABRIEL, U., WEISS, C., ERBEN, P., MARTINI, T. & BOLENZ, C. 2014. IL1RN and KRT13 Expression in Bladder Cancer: Association with Pathologic Characteristics and Smoking Status. *Adv Urol,* 2014**,** 184602.

WRIGHT, D. H., STONE, J. A., CRUMLEY, T. M., WENNING, L., ZHENG, W., YAN, K., YANG, A. Y., SUN, L., CILISSEN, C., RAMAEL, S., HERMANOWSKI-VOSATKA, A., LANGDON, R. B., GOTTESDIENER, K. M., WAGNER, J. A. & LAI, E. 2013. Pharmacokinetic-pharmacodynamic studies of the 11beta-hydroxysteroid dehydrogenase type 1 inhibitor MK-0916 in healthy subjects. *Br J Clin Pharmacol,* 76**,** 917-31.

WU, B., GUO, B. M., KANG, J., DENG, X. Z., FAN, Y. B., ZHANG, X. P. & AI, K. X. 2016a. PPM1D exerts its oncogenic properties in human pancreatic cancer through multiple mechanisms. *Apoptosis,* 21**,** 365-78.

WU, G. & VANCE, D. E. 2010. Choline kinase and its function. *Biochem Cell Biol,* 88**,** 559-64.

WU, X., LIU, T., FANG, O., LEACH, L. J., HU, X. & LUO, Z. 2014. miR-194 suppresses metastasis of non-small cell lung cancer through regulating expression of BMP1 and p27(kip1). *Oncogene,* 33**,** 1506-14.

WU, Y., GU, Y., GUO, S., DAI, Q. & ZHANG, W. 2016b. Expressing Status and Correlation of ARID1A and Histone H2B on Breast Cancer. *Biomed Res Int,* 2016**,** 7593787.

WU Z, I. R., GENTLEMAN R, MURILLO FM, SPENCER F 2004. A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association***,** 909-917.

XIE, M., SUN, M., ZHU, Y. N., XIA, R., LIU, Y. W., DING, J., MA, H. W., HE, X. Z., ZHANG, Z. H., LIU, Z. J., LIU, X. H. & DE, W. 2015. Long noncoding RNA HOXA-AS2 promotes gastric cancer proliferation by epigenetically silencing P21/PLK3/DDIT3 expression. *Oncotarget*.

XING, C., XIE, H., ZHOU, L., ZHOU, W., ZHANG, W., DING, S., WEI, B., YU, X., SU, R. & ZHENG, S. 2012. Cyclin-dependent kinase inhibitor 3 is overexpressed in hepatocellular carcinoma and promotes tumor cell proliferation. *Biochem Biophys Res Commun,* 420**,** 29-35.

XU, J., WENG, Z., ARUMUGAM, A., TANG, X., CHAUDHARY, S. C., LI, C., CHRISTIANO, A. M., ELMETS, C. A., BICKERS, D. R. & ATHAR, M. 2014. Hair follicle disruption facilitates pathogenesis to UVB-induced cutaneous inflammation and basal cell carcinoma development in Ptch(+/-) mice. *Am J Pathol,* 184**,** 1529-40.

XU, L., TANG, H., CHEN, D. W., EL-NAGGAR, A. K., WEI, P. & STURGIS, E. M. 2015. Genome-wide association study identifies common genetic variants associated with salivary gland carcinoma and its subtypes. *Cancer,* 121**,** 2367-74.

XU, X. L., LI, B., SUN, X. L., LI, L. Q., REN, R. J., GAO, F. & JONAS, J. B. 2008. Eyelid neoplasms in the Beijing Tongren Eye Centre between 1997 and 2006. *Ophthalmic Surg Lasers Imaging,* 39**,** 367-72.

XU, Y., XIA, Q., RAO, Q., SHI, S., SHI, Q., MA, H., LU, Z., CHEN, H. & ZHOU, X. 2016. DCN deficiency promotes renal cell carcinoma growth and metastasis through downregulation of P21 and E-cadherin. *Tumour Biol,* 37**,** 5171-83.

XUE, M., CHEN, W. & LI, X. 2015. Urothelial cancer associated 1: a long noncoding RNA with a crucial role in cancer. *J Cancer Res Clin Oncol*.

YAJIMA, I., KUMASAKA, M. Y., YAMANOSHITA, O., ZOU, C., LI, X., OHGAMI, N. & KATO, M. 2014. GNG2 inhibits invasion of human malignant melanoma cells with decreased FAK activity. *Am J Cancer Res,* 4**,** 182-8.

YAKOVLEV, A. G., DI GIOVANNI, S., WANG, G., LIU, W., STOICA, B. & FADEN, A. I. 2004. BOK and NOXA are essential mediators of p53-dependent apoptosis. *J Biol Chem,* 279**,** 28367-74.

YAMAGIWA, K. A. I., K 1916. Über die Künstliche Erzeugung von Carcinom. *Gann,* 10**,** 21-33.

YAN, H., WANG, S., YU, H., ZHU, J. & CHEN, C. 2013. Molecular pathways and functional analysis of miRNA expression associated with paclitaxel-induced apoptosis in hepatocellular carcinoma cells. *Pharmacology,* 92**,** 167-74.

YANG, C. & SUN, J. J. 2015. Mechanistic studies of cyclin-dependent kinase inhibitor 3 (CDKN3) in colorectal cancer. *Asian Pac J Cancer Prev,* 16**,** 965-70.

YANG, J. J., CHENG, C., YANG, W., PEI, D., CAO, X., FAN, Y., POUNDS, S. B., NEALE, G., TREVINO, L. R., FRENCH, D., CAMPANA, D., DOWNING, J. R., EVANS, W. E., PUI, C. H., DEVIDAS, M., BOWMAN, W. P., CAMITTA, B. M., WILLMAN, C. L., DAVIES, S. M., BOROWITZ, M. J., CARROLL, W. L., HUNGER, S. P. & RELLING, M. V. 2009. Genome-wide interrogation of germline genetic variation associated with treatment response in childhood acute lymphoblastic leukemia. *Jama,* 301**,** 393-403.

YANG, N. Y., PASQUALE, E. B., OWEN, L. B. & ETHELL, I. M. 2006. The EphB4 receptor-tyrosine kinase promotes the migration of melanoma cells through Rho-mediated actin cytoskeleton reorganization. *J Biol Chem,* 281**,** 32574-86.

YAO, D. M., ZHOU, J. D., ZHANG, Y. Y., YANG, L., WEN, X. M., YANG, J., GUO, H., CHEN, Q., LIN, J. & QIAN, J. 2015. GPX3 promoter is methylated in chronic myeloid leukemia. *Int J Clin Exp Pathol,* 8**,** 6450-7.

YAUCH, R. L., DIJKGRAAF, G. J., ALICKE, B., JANUARIO, T., AHN, C. P., HOLCOMB, T., PUJARA, K., STINSON, J., CALLAHAN, C. A., TANG, T., BAZAN, J. F., KAN, Z., SESHAGIRI, S., HANN, C. L., GOULD, S. E., LOW,

J. A., RUDIN, C. M. & DE SAUVAGE, F. J. 2009. Smoothened mutation confers resistance to a Hedgehog pathway inhibitor in medulloblastoma. *Science,* 326**,** 572-4.

YAUCH, R. L., GOULD, S. E., SCALES, S. J., TANG, T., TIAN, H., AHN, C. P., MARSHALL, D., FU, L., JANUARIO, T., KALLOP, D., NANNINI-PEPE, M., KOTKOW, K., MARSTERS, J. C., RUBIN, L. L. & DE SAUVAGE, F. J. 2008. A paracrine requirement for hedgehog signalling in cancer. *Nature,* 455**,** 406-10.

YEH, C. T., LU, S. C., CHAO, C. H. & CHAO, M. L. 2003. Abolishment of the interaction between cyclin-dependent kinase 2 and Cdk-associated protein phosphatase by a truncated KAP mutant. *Biochem Biophys Res Commun,* 305**,** 311-4.

YOU, J., FANG, N., GU, J., ZHANG, Y., LI, X., ZU, L. & ZHOU, Q. 2014. Noncoding RNA small nucleolar RNA host gene 1 promote cell proliferation in nonsmall cell lung cancer. *Indian J Cancer,* 51 Suppl 3**,** e99-e102.

YU, F., BRACKEN, C. P., PILLMAN, K. A., LAWRENCE, D. M., GOODALL, G. J., CALLEN, D. F. & NEILSEN, P. M. 2015a. p53 Represses the Oncogenic Sno-MiR-28 Derived from a SnoRNA. *PLoS One,* 10**,** e0129190.

YU, J., LI, A., HONG, S. M., HRUBAN, R. H. & GOGGINS, M. 2012. MicroRNA Alterations of Pancreatic Intraepithelial Neoplasms (PanINs). *Clin Cancer Res,* 18**,** 981-92.

YU, X., LI, Y., CHEN, S. W., SHI, Y. & XU, F. 2015b. Differential expression of glypican-3 (GPC3) in lung squamous cell carcinoma and lung adenocarcinoma and its clinical significance. *Genet Mol Res,* 14**,** 10185-92.

YU, Y. N., YIP, G. W., TAN, P. H., THIKE, A. A., MATSUMOTO, K., TSUJIMOTO, M. & BAY, B. H. 2010. Y-box binding protein 1 is up-regulated in proliferative breast cancer and its inhibition deregulates the cell cycle. *Int J Oncol,* 37**,** 483-92.

YUAN, Z. Y., DAI, T., WANG, S. S., PENG, R. J., LI, X. H., QIN, T., SONG, L. B. & WANG, X. 2014. Overexpression of ETV4 protein in triple-negative breast cancer is associated with a higher risk of distant metastasis. *Onco Targets Ther,* 7**,** 1733-42.

YUEN, T. J., JOHNSON, K. R., MIRON, V. E., ZHAO, C., QUANDT, J., HARRISINGH, M. C., SWIRE, M., WILLIAMS, A., MCFARLAND, H. F., FRANKLIN, R. J. & FFRENCH-CONSTANT, C. 2013. Identification of endothelin 2 as an inflammatory factor that promotes central nervous system remyelination. *Brain,* 136**,** 1035-47.

YUVA-AYDEMIR, Y., SIMKIN, A., GASCON, E. & GAO, F. B. 2011. MicroRNA-9: functional evolution of a conserved small regulatory RNA. *RNA Biol,* 8**,** 557-64.

ZEEUWEN, P. L., ISHIDA-YAMAMOTO, A., VAN VLIJMEN-WILLEMS, I. M., CHENG, T., BERGERS, M., IIZUKA, H. & SCHALKWIJK, J. 2007. Colocalization of cystatin M/E and cathepsin V in lamellar granules and corneodesmosomes suggests a functional role in epidermal differentiation. *J Invest Dermatol,* 127**,** 120-8.

ZEILSTRA, J., JOOSTEN, S. P., WENSVEEN, F. M., DESSING, M. C., SCHUTZE, D. M., ELDERING, E., SPAARGAREN, M. & PALS, S. T. 2011. WNT signaling controls expression of pro-apoptotic BOK and BAX in intestinal cancer. *Biochem Biophys Res Commun,* 406**,** 1-6.

ZETTL, M. & WAY, M. 2002. The WH1 and EVH1 domains of WASP and Ena/VASP family members bind distinct sequence motifs. *Curr Biol,* 12**,** 1617-22.

ZHA, H., SUN, H., LI, X., DUAN, L., LI, A., GU, Y., ZENG, Z., ZHAO, J., XIE, J., YUAN, S., LI, H. & ZHOU, L. 2016. S100A8 facilitates the migration of colorectal cancer cells through regulating macrophages in the inflammatory microenvironment. *Oncol Rep,* 36**,** 279-90.

ZHANG, B., JIA, W. H., MATSUDA, K., KWEON, S. S., MATSUO, K., XIANG, Y. B., SHIN, A., JEE, S. H., KIM, D. H., CAI, Q., LONG, J., SHI, J., WEN, W., YANG, G., ZHANG, Y., LI, C., LI, B., GUO, Y., REN, Z., JI, B. T., PAN, Z. Z., TAKAHASHI, A., SHIN, M. H., MATSUDA, F., GAO, Y. T., OH, J. H., KIM, S., AHN, Y. O., CHAN, A. T., CHANG-CLAUDE, J., SLATTERY, M. L., GRUBER, S. B., SCHUMACHER, F. R., STENZEL, S. L., CASEY, G., KIM, H. R., JEONG, J. Y., PARK, J. W., LI, H. L., HOSONO, S., CHO, S. H., KUBO, M., SHU, X. O., ZENG, Y. X. & ZHENG, W. 2014. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat Genet,* 46**,** 533-42.

ZHANG, G., GU, D., ZHAO, Q., CHU, H., XU, Z., WANG, M., TANG, C., WU, D., TONG, N., GONG, W., ZHOU, J., XU, Y., ZHANG, Z. & CHEN, J. 2015a. Genetic variation in C12orf51 is associated with prognosis of intestinal-type gastric cancer in a Chinese population. *Biomed Pharmacother,* 69**,** 133-8.

ZHANG, L., MEI, Y., FU, N. Y., GUAN, L., XIE, W., LIU, H. H., YU, C. D., YIN, Z., YU, V. C. & YOU, H. 2012. TRIM39 regulates cell cycle progression and DNA damage responses via stabilizing p21. *Proc Natl Acad Sci U S A,* 109**,** 20937-42.

ZHANG, L. P., LI, W. J., ZHU, Y. F., HUANG, S. Y., FANG, S. Y., SHEN, L. & GAO, Y. L. 2015b. CDKN3 knockdown reduces cell proliferation, invasion and promotes apoptosis in human ovarian cancer. *Int J Clin Exp Pathol,* 8**,** 4535-44.

ZHANG, L. Y., WANG, X. L., SUN, D. X., LIU, X. X., HU, X. Y. & KONG, F. 2008. Regulation of zinc transporters by dietary flaxseed lignan in human breast cancer xenografts. *Mol Biol Rep,* 35**,** 595-600.

ZHANG, W., YAO, J. L., DONG, S. C., HOU, F. Q. & SHI, H. P. 2015c. SLPI knockdown induced pancreatic ductal adenocarcinoma cells proliferation and invasion. *Cancer Cell Int,* 15**,** 37.

ZHANG, X., ZHOU, J., REEDERS, S. T. & TRYGGVASON, K. 1996. Structure of the human type IV collagen COL4A6 gene, which is mutated in Alport syndrome-associated leiomyomatosis. *Genomics,* 33**,** 473-9.

ZHANG, Y., DEWITT, D. L., MCNEELY, T. B., WAHL, S. M. & WAHL, L. M. 1997. Secretory leukocyte protease inhibitor suppresses the production of monocyte prostaglandin H synthase-2, prostaglandin E2, and matrix metalloproteinases. *J Clin Invest,* 99**,** 894-900.

ZHAO, B., SCHLESIGER, C., MASUCCI, M. G. & LINDSTEN, K. 2009. The ubiquitin specific protease 4 (USP4) is a new player in the Wnt signalling pathway. *J Cell Mol Med,* 13**,** 1886-95.

ZHAO, H., LAI, X., XU, X., SUI, K., BU, X., MA, W., LI, D., GUO, K., XU, J., YAO, L., LI, W. & SU, J. 2015a. Histochemical analysis of testis specific gene 13 in human normal and malignant tissues. *Cell Tissue Res,* 362**,** 653-63.

ZHAO, J., CHEN, J., LIN, H., JIN, R., LIU, J., LIU, X., MENG, N. & CAI, X. 2015b. The Clinicopathologic Significance of BAF250a (ARID1A) Expression in Hepatocellular Carcinoma. *Pathol Oncol Res*.

ZHAO, L., WEI, Y., SONG, A. & LI, Y. 2016. Association study between genome-wide significant variants of vitamin B12 metabolism and gastric cancer in a han Chinese population. *IUBMB Life,* 68**,** 303-10.

ZHOU, C., WU, Y. L., CHEN, G., FENG, J., LIU, X. Q., WANG, C., ZHANG, S., WANG, J., ZHOU, S., REN, S., LU, S., ZHANG, L., HU, C., LUO, Y., CHEN, L., YE, M., HUANG, J., ZHI, X., ZHANG, Y., XIU, Q., MA, J. & YOU, C. 2011. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol,* 12**,** 735-42.

ZHOU, P., CHEN, W. G. & LI, X. W. 2015. MicroRNA-143 acts as a tumor suppressor by targeting hexokinase 2 in human prostate cancer. *Am J Cancer Res,* 5**,** 2056-63.

ZHOU, Y., ZHU, J., ZHANG, Y., JIANG, J. & JIA, M. 2013. An inflammatory myofibroblastic tumour of the breast with ALK overexpression. *BMJ Case Rep,* 2013.

ZHU, C., XIAO, F. & LIN, W. 2015. EFTUD2 on innate immunity. *Oncotarget*.

ZHU, C. L. & GAO, G. S. 2014. miR-200a overexpression in advanced ovarian carcinomas as a prognostic indicator. *Asian Pac J Cancer Prev,* 15**,** 8595-601.

ZOU, Y., LIN, X., FAN, D., CHEN, X., YANG, Z., ZHENG, X., LIU, X., WU, X. & LAN, P. 2015. [Expression and clinical significance of G protein-coupled receptor 31 in colorectal cancer tissue]. *Zhonghua Wei Chang Wai Ke Za Zhi,* 18**,** 935-40.

ZUO, Y. & SONG, Y. 2013. Detection and analysis of the methylation status of PTCH1 gene involved in the hedgehog signaling pathway in a human gastric cancer cell line. *Exp Ther Med,* 6**,** 1365-1368.